Growing through Competition: The Reduction of Entry Barriers among Chinese Manufacturing Firms *

Helu Jiang (Shanghai University of Finance and Economics)†
Yu Zheng (Queen Mary University of London and CEPR)‡
Lijun Zhu (Peking University)§

April 30, 2021

Abstract

Exploiting the gradualism of the Chinese economic reforms and cross-sectional variations in entry rates, we show empirical evidence from firm-level data that industries with higher entry rates achieve higher growth and a more competitive market structure in subsequent years. We then embed firm entry into a model of endogenous productivity and market structure with heterogeneous firms and sectors, and calibrate it to the Chinese manufacturing sector in 2004-7. We find the positive impact of entry on growth is achieved primarily through a *pro-competitive effect*, whereby entry induces endogenously a larger fraction of industries to be more competitive in the economy. We quantify the contribution on growth from the reduction of entry barriers associated with the state-owned enterprise reforms in the late 1990s and early 2000s and find it explains more than 40% of the aggregate growth differentials of the manufacturing sector between 1991-5 and 2004-7. More generally, we highlight the critical role of reducing entry barriers in promoting competition and growth in developing countries.

JEL classification: D22, D43, O11, O30, O47

Keywords: Firm Entry; Endogenous Growth; Firm Dynamics; Entry Barriers

---

*We gratefully acknowledge helpful comments from Sina Ates, Paco Buera, Chad Jones, Pete Klenow, and Justin Yifu Lin as well as the seminar and conference participants at Bristol, UCL, Essex, Jinan University, Fudan University and the 7th NSE International Conference. All remaining errors are ours.

†Institute for Advanced Research, Shanghai University of Finance and Economics. 111 Wuchuan Road, Shanghai, China, 200433. Email: jianghelu@sufe.edu.cn.

‡Corresponding author: School of Economics and Finance, Queen Mary University of London. Mile End Road, E1 4NS London, UK. Email: yu.zheng@qmul.ac.uk.

§Institute of New Structural Economics, Peking University. No. 5 Yiheyuan Road, Haidian District, Beijing, China 100871. Email: lijunzhu@nsd.pku.edu.cn.

1 Introduction

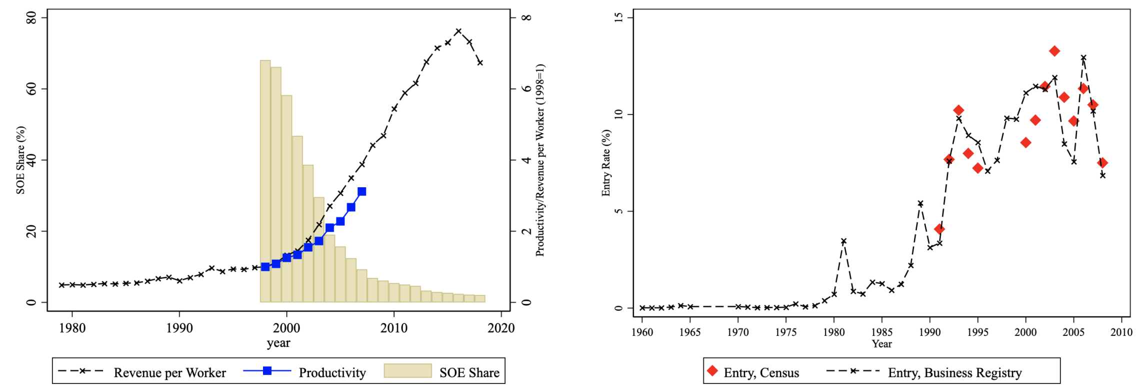
The idea that monopoly stifles growth while competition promotes it is hardly new. Writing on the rise of the western world during 1500-1700, North and Thomas (1973) ascribe the stagnation of France to the industrial regulation and the guild system that granted monopoly to insiders and restricted entry of outsiders; In England, in contrast, new rules like the Statute of Monopolies introduced in the early 17th century stroke down monopolistic privileges and barriers to entry, which previously circumscribed profitable opportunities in trade and commerce, and eventually set the stage for the industrial revolution. This historical view is echoed by many observers of China's reforms and industrialization since the late 1970s, when state monopoly was cut back, private firm entry permitted, and state-owned enterprises privatized. Meanwhile the labor productivity of the industrial sector increased from 2.2% in the 1980s, to 8.47% in the 1990s and 15.27% in the 2000s (Figure 1.1(a)). The force of incentives and competition released in the process is deemed to be a critical pillar underpinning the success of the reform (McMillan and Naughton, 1992; Groves et al., 1994; Qian, 2002; Brandt et al., 2008; Zhu, 2012).

In this paper, we revisit this view and study the pro-competition and pro-growth effects of reduced entry barriers to the Chinese manufacturing sector during this economic transition to a market economy. Like Brandt et al. (2012) and Brandt et al. (2020), we exploit large-scale firm-level data such as the Annual Survey of Industrial Enterprises (ASIE), Industrial Census and Business Registry Records to establish facts and discipline our theory. Different than previous work, we propose a theory that features endogenous productivity, entry and market structure with heterogeneous firms and sectors to assess quantitatively the role entry plays in achieving the higher productivity growth in the recent decades.

Our starting point is the recognition that historically entry barriers were reduced across different industries in the industrial sector in a staggered fashion throughout the 1980s, 1990s and well into the 2000s (Figure 1.1(b)), and therefore provide us with a dispersion of entry rates across four-digit industries in as late as 2004 (Figure 1.2(a)).[1] While there exist large differences in entry rates and state presence across 2-digit industries (e.g. comparing transportation equipment to furniture manufacturing), even within a 2-digit industry, we observe a sizeable dispersion of entry rates. For example, within food manufacturing, soy sauce and vinegar manufacturing has large state presence and low entry

[1] In Appendix B.2, we confirm the broad trends of declining SOE share and rising output per worker are experienced in almost all industries within the manufacturing sector, but the changes occurred at different speed.

Figure 1.1: Aggregate Productivity, SOE Share and Entry Rate in the Chinese Industrial Sector, 1960-2014



(a) Productivity and SOE Share

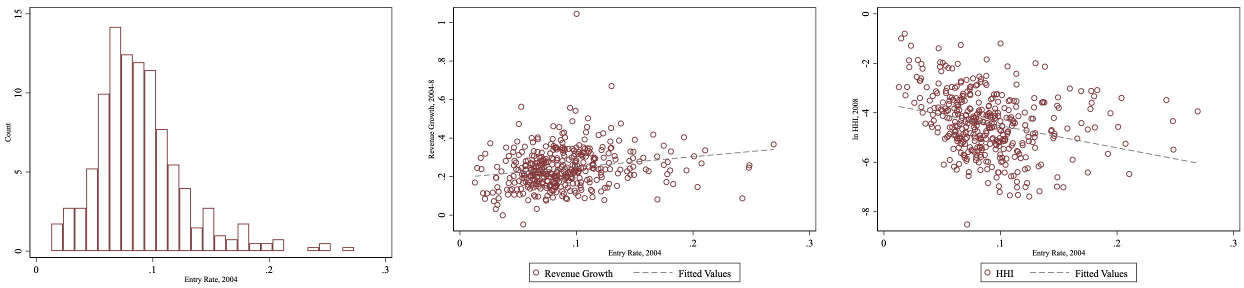
(b) Aggregate Entry Rate

Note: This figure shows the labor productivity (real revenue per worker) and the share of SOE based on the NBS data together with our employment-weighted productivity estimated from the ASIE (Panel (a)) and aggregate entry rates constructed from the Industrial Census and the Business Registry Records (Panel (b)). The construction of the series is detailed in Appendix A.4.

rates, whereas frozen food manufacturing has a much higher entry rate in 2004. This cross-sectional variation of entry induced by the piecemeal reform allows us to investigate how entry affects growth and market structure across industries. Panels (b) and (c) of Figure 1.2 show raw correlations of four-digit industry-level entry rates in 2004 with the industry-level real revenue growth over 2004-8 and industry concentration measured by the Herfindahl-Hirschman Index (HHI) in 2008. Industries whose entry barrier is lowered by 2004 and hence having a high entry rate in 2004 tend to grow faster and become more competitive. In the empirical section of the paper, we will make these statements more precise, properly controlling for 2-digit industry fixed effects and using additional data from the ASIE panel of manufacturing firms. We show empirical evidence that as entry barriers are gradually lifted and entry occurs, tighter competition ensues and productivity growth accelerates. We then propose a macro model to explain these facts.

To have a meaningful discussion on the impact of entry on the level of competition, we deviate from prior literature which has viewed data through the lens of a Hopenhayn (1992) or Melitz (2003) model. Instead, we interpret the Chinese data using a model of endogenous productivity in the spirit of the step-by-step innovation model of Aghion et al. (2001), extended to include heterogeneous firms and sectors. It is noteworthy however that, in the Chinese context, productivity growth does not necessarily come from the

Figure 1.2: Entry, HHI and Revenue Growth in the Chinese Industrial Sector, 2004-2008 Growth



(a) Distribution of Entry Rate (b) Entry and Revenue Growth (c) Entry and HHI

Note: This figure shows the distribution of entry rates across 4-digit industries in 2004 (Panel (a)), the scatter plot of these entry rates against the growth of industry total revenue from 2004 to 2008 (Panel (b)) and the industry concentration measure, HHI, in 2008 (Panel (c)), based on data from Industrial Census 2004 and 2008. The construction of the series is detailed in Appendix A.4.

narrow definition of technological innovation. Any costly activity to improve organizational efficiency, secure a stable supply chain, increase brand awareness, or sell to a new geographical or demographic market segment can be viewed as an avenue for improving productivity of the firm and thus falls in the realm of what we call "innovation" in this paper.[2] The economy consists of two productive sectors with differential entry barriers, and within each sector is a continuum of symmetric industries. There is a quality ladder which firms compete to climb. In each industry, there are two incumbent firms; the firm that is ahead on the quality ladder is the market leader, and the one lagging behind is the market follower. The leader and follower, which produce goods that are imperfect substitutes, engage in Bertrand competition. The relative market share of the leader to the follower increases in the distance on the quality ladder between the two, while in a neck-and-neck industry where the distance between the two incumbents is zero, the market shares of the two are equal. Furthermore, firm's revenue is a logistic function of its distance to the opponent. As a result, firms in industries where the leader-follower quality gap is smaller and therefore more competitive, have stronger incentives to innovate and escape competition, while those in industries with bigger gaps innovate significantly less.

There is a potential entrant in each industry at any point in time; successful entrants replace followers in industries with gaps, and randomly replace incumbent firms in neck-

[2] Ates and Saffie (2020) and Peters (2020) employ a similar approach to interpret models of innovation as models of endogenous productivity.

and-neck industries. Firms can be one of two types: The high (low) type has lower (higher) cost of innovation. We assume that firms enter as a high type, transit into a low type at a random rate over time and the low type is an absorbing type. This assumption is motivated by the fact that older firms grow significantly more slowly than younger ones in the Chinese Manufacturing sector. We focus on the stationary equilibrium and show that the aggregate growth rate in the stationary equilibrium is determined by all firms' innovation efforts in neck-and-neck industries and leaders' innovation efforts in all other industries in the two sectors. We calibrate this model to the Chinese manufacturing sector in 2004-7, where staggered reform leaves a wide dispersion of entry rates across industries in those years.

The main advantage of adopting such a step-by-step endogenous growth framework, apart from our motivation to understand the acceleration of growth over long periods of time in China, is that the model gives rise to an endogenous distribution of market structure across industries in the equilibrium. This is in contrast to models of perfect competition or monopolistic competition, where entry does not impact the market structure per se, and therefore entry and competition are synonyms. In those models, entry relies on a selection effect alone to achieve productivity gain. Namely, more productive entrants replace less productive incumbents. In our model, we however identify four channels through which entry affects growth, one of which, the Schumpeterian effect, represents actually a negative impact on growth following entry. Of the four channels, the *replacement effect*, namely highly productive entrants replacing less productive incumbents and therefore increasing the prevalence of the more productive type in the economy, resembles the selection effect in the prior literature. However, the *pro-competitive effect*, whereby entry increases the fraction of relatively more competitive industries, is new. And we show that this new effect is the dominant channel through which high-entry industries realize high-growth in 2004-7.

Using the calibrated model, we conduct a counterfactual analysis to quantify the amount of aggregate growth in the Chinese manufacturing sector over 2004-7 that is generated by the increase in entry which is associated with the SOE reforms in the late 1990s and early 2000s. To isolate the amount of entry which is induced by the SOE reforms, we provide more empirical evidence and estimate the elasticity of entry with respect to the presence of SOEs in an industry from our census sample. Using the counterfactual prediction of the level of entry based on the difference in the SOE presence between 1995 and 2004, we construct the targeted counterfactual entry rate in the pre-reform year of 1995 and recalibrate the model. We find that the reform-induced entry accounts for more than 40% of the aggregate growth differentials experienced by the manufacturing sector between 1991-5 and 2004-7. Of the growth differential due to the reform-induced entry, 39% stems from the replacement effect and 57% stems from the pro-competitive effect. These results underscore once again the importance of adopting a model which permits the competitiveness of industries to endogenously respond to entry.

The paper is related to three strands of literature. The first strand of literature investigates the mechanisms behind China's economic growth. This includes, but is not limited to, the expansion of the non-state sector (Zhu, 2012; Hsieh and Song, 2015), the reduction of entry barriers (Brandt et al., 2012, 2020); the improved allocation of capital (Song et al., 2011); and more generally the reduction in inefficiencies in output and factor markets (Hsieh and Klenow, 2009; Cheremukhin et al., 2017). We contribute to this literature by adopting an endogenous growth model with endogenous market structure to study the relationship between entry, competition and growth in the Chinese context. The second strand of literature our paper relates to is Schumpeterian growth models with step by step innovation (Aghion et al., 2001, 2005a). Akcigit and Ates (2019) extend the model to incorporate entry to study the declining business dynamism in the United States. We build on this class of models by introducing entry, heterogeneous firms and sectors to adapt to the Chinese context, and assess quantitatively the impact of reducing entry barriers in the previously state-dominated sectors.

The third strand of literature examines more broadly the role of entry barrier in explaining economic growth in developing countries or the lack thereof and the economic inequality in development (Parente and Prescott, 1999; Aghion et al., 2005b; Herrendorf and Teixeira, 2011; Asturias et al., 2019). In particular, Asturias et al. (2019) study the role of entry in a Hopenhayn style growth model, in which the productivity distribution from which entrants draw grows at an exogenous rate. In contrast, all cohorts of firms in our model are equally productive in the sense that they all start from a high type gradually transiting into a low type at the same rate. The pro-competitive effect of firm entry on aggregate growth is achieved by industries evolving into a more competitive market structure, providing incentives for incumbent firms as well as entrants to pursue growth. More recently, Peters (2020) examines the costs of entry and of expansion in explaining productivity differences between the US and Indonesia in a model of endogenous markup and productivity, where creative destruction by entrants can lower markup and promote competition, however their effects on growth rates are muted. The Chinese experience we study is unique in the magnitude and duration of the productivity growth realized and offers support to the view that policies that reduce entry barriers and unleash competition are effective policies to deliver growth.

The rest of the paper is organized as follows. In Section 2, we describe our data and present empirical evidence for the pro-competitive and growth-enhancing effects of entry. In Section 3, we present the two-sector model with endogenous productivity and entry and heterogeneous firms and sectors. In Section 4, we calibrate the model to the Chinese manufacturing sector in 2004-7 and provide a growth rate decomposition to highlight the various channels through which entry affects growth. In Section 5, we provide more evidence that relates the entry rates to the presence of SOEs and assess counterfactually the contribution to growth from reducing the entry barriers during the SOE reforms in the late 1990s and early 2000s. Conclusion follows.

2 Empirical Motivation and Evidence

2.1 Data, Sample Construction, and Summary Statistics

Our main data sources are the Chinese Industrial Census 1995, 2004 and 2008 as well as the Annual Surveys of Industrial Enterprises (ASIE) from 1998 to 2007, both conducted by the National Bureau of Statistics. We use the Industrial Census, also known as the Economic Census, to compute the entry rates and measures of competition by industry, as the census includes all operating firms in a given year. The ASIE, on the other hand, is a panel of "above scale" industrial firms, i.e. firms with annual sales above 5 million RMB, and we use it to estimate firm-level productivity which requires a panel.[3] We summarize here the construction of variables and the analysis samples and present the summary statistics from the samples. More details can be found in Appendix A.

Entry Entry rates are defined at 4-digit CIC industry level. In order to alleviate the issue of reporting errors, we use 2-year averages to calculate entry rates from the census sample. For example, from the 2004 Census sample, we define the entry rate to be the ratio

[3]It's well known that the above-scale firms capture around 90% of the total industrial output and 70% of the total industrial employment. By comparing the ASIE to the Industrial Census, Brandt et al. (2012) reach the conclusion that in 2004 below-scale firms employed 28.8% of the industrial workforce, produced 9.9% of output and 2.5% of exports. Comparison with the 1995 Census yields similar results.

of the average number of firms established in 2003 and 2004 to the number of existing firms in 2004 using employment as weight within each 4-digit industry.

Competition We use HHI at 4-digit CIC industry level as the measure of competition in a narrowly defined industry. HHI is defined as the sum of squared revenue shares of all firms within an industry.

Productivity Given our endogenous growth framework, we modify the two-stage control function approach advocated by De Loecker and Warzynski (2012) and Levinsohn and Petrin (2003) to estimate firm-level productivity. More specifically, we estimate with GMM a value-added Cobb-Douglas production function at the 2-digit CIC industry level, using material demand as the proxy for the unobserved persistent productivity shock and assuming that the persistent productivity shocks follow a random walk with a drift, which is itself a function of the current period's capital and last period's HHI. We also allow an indicator of state ownership to affect both the demand for materials and the random walk process. The procedure is detailed in Appendix A.2.

Table 2.1: Summary Statistics, Industrial Census and ASIE

	Census Sample		
	1995	2004	2008
Age	12.30	6.57	6.40
	(12.93)	(8.22)	(6.95)
SOE (%)	89.34	12.03	4.21
	(30.86)	(32.53)	(20.07)
Revenue (thousand)	10,865.85	14,619.28	23,823.97
	(126,402.74)	(259,951.13)	(443,324.12)
Employment (persons)	164.25	65.57	58.21
	(860.51)	(344.67)	(372.90)
Number of firms per industry-year		82,695.58	
		(56,069.66)	
Number of industries		29	
Number of industry-year observations		87	
	ASIE Sample, Selected Years		
	1998	2002	2007
Age	15.80	14.13	10.53
	(13.80)	(12.51)	(9.81)
SOE (%)	28.77	13.37	3.44
	(45.27)	(34.03)	(18.23)
Value added (thousand)	7,827.91	10,194.01	14,804.44
	(17,964.91)	(21,845.76)	(31,296.14)
Employment (persons)	273.02	218.06	161.50
	(469.37)	(328.77)	(246.32)
Productivity	1.00	1.42	2.61
	(0.87)	(1.16)	(2.24)
Number of firms per industry-year		7,301.66	
		(4,427.27)	
Number of industries		26	
Number of industry-year observations		260	

Note: This table reports the mean and standard deviations (in parentheses) of the key variables for the census and the ASIE sample.

We restrict our sample to the manufacturing sector in both the ASIE and census samples, that is all 4-digit CIC codes between 1300 and 4400. The sample selection process is detailed in Appendix A.3. Our census sample is three repeated cross-sections of firms in 29 two-digit industries. Our ASIE sample is an unbalanced panel of firms in 26 two-digit industries from 1998 to 2007. Table 2.1 provides the summary statistics of the key economic variables from the census sample and selected years from the ASIE sample. Revenue and value added are in thousands of 1998 Chinese *yuan*. Due to the sampling difference, the census sample has more young and small firms than the ASIE sample. Due to the difference in the definition of SOE, where in the census sample it is based on registered type of business (Brandt et al., 2012) and in the ASIE sample it is based on equity ownership (Hsieh and Song, 2015), the level of SOE shares differs across the two samples. In both samples, the broad trends are however similar, that is firms on average become younger, less state owned, and bigger in terms of output and smaller in terms of employment over time.

2.2 Relationship between Entry, Growth, and Competition

In this section, we present more rigorous empirical analysis of the pattern of correlations presented in Figure 1.2 in the introduction, using the Census sample of the entire universe of firms in the manufacturing sector as well as the analysis panel of above scale manufacturing firms from the ASIE.

Firstly, we construct from the Census sample, at 4-digit industry level, entry rates in 2004, total industry revenue growth from 2004 to 2008 and (log) HHI in 2008. When we regress the revenue growth on entry rate, controlling for 2-digit industry fixed effects and initial 4-digit industry attributes such as number of firms, employment and revenue, and cluster the standard errors at the 2-digit industry level, we find significant positive effects of entry on revenue growth (Column [1]-[3] in Table 2.2). For one percentage point increase in entry rate, the industry revenue growth accelerates by roughly half a percentage point. At the same time, industries with higher entry in 2004 tend to evolve into a market structure that is less concentrated as measured by HHI in 2008. Column [4]-[6] of the same table show that one percentage point increase in entry rate tends to reduce HHI by about 4%. The industry-level evidence points to higher growth and tougher competition which concur following entry.[4, 5]

Now we turn to direct firm-level evidence from the more restrictive ASIE sample. Entry naturally brings in young firms, which typically experience higher growth at the ex-

[4]In Appendix B.1, we also show the significantly negative correlation between HHI and revenue growth at the industry level using the same 2004 and 2008 Census data.

[5]In Appendix B.2, we show in Table B.4 that such relationships between entry, revenue growth, and competition are mainly driven by the entry of private-owned enterprises as opposed to state-owned or foreign owned enterprises.

Table 2.2: Average Industry Entry Rate, Industry Real Revenue Growth and Industry HHI

	Real Revenue Growth			log HHI		
	(1)	(2)	(3)	(4)	(5)	(6)
average entry rate	0.805***	0.511***	0.473***	-6.046***	-4.533***	-3.757**
	(5.51)	(3.71)	(3.77)	(-3.94)	(-3.07)	(-2.47)
2004 number of firms (million)	0.109	-0.788	-0.948	-123.2***	-118.6***	-115.3***
	(0.16)	(-1.23)	(-1.43)	(-6.22)	(-6.36)	(-6.09)
2004 log industry employment		-0.0473***	-0.0229		0.243***	-0.256
		(-5.89)	(-1.19)		(2.81)	(-1.66)
2004 log industry revenue			-0.0207			0.424***
			(-1.48)			(3.80)
R^2	0.183	0.238	0.245	0.517	0.528	0.548
Observations	400	400	400	400	400	400

t statistics in parentheses

2-digit CIC industry fixed effects controlled; standard errors clustered at 2-digit industry level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

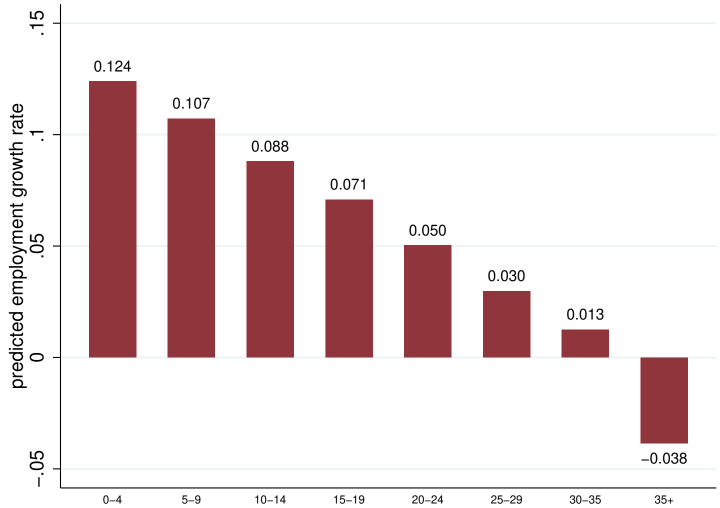
Note: This table reports the results regressing the annual industry-level revenue growth rate (column [1] to [3]) and industry-level log HHI (column [4] to [6]) on entry rates, controlling for the initial industry characteristics such as the number of firms, industry employment and revenue, as well as 2-digit CIC industry fixed effects, with standard errors clustered at 2-digit CIC industry level. Entry rate is defined as the ratio of the average number of firms established in 2003 and 2004 to the total number of incumbent firms in 2004, weighted by employment. Total real revenue is calculated by industry in 2004 and 2008 and growth rate is calculated as the average annual growth rate within the time period. HHI index is calculated by industry in 2008.

pansion stage. This is a well-known fact, which has been established for the US by for example Haltiwanger et al. (2016). It is also true for our sample of Chinese manufacturing firms, as illustrated by Figure 2.1. Using employment as the measure of firm size, we regress firm-level annual employment growth on firm's age, controlling for 4-digit industry fixed effect and firm-level characteristics such as employment, capital, and export status, clustering the standard errors at the 2-digit industry level using the 2005-7 ASIE panel. Then we plot the average predicted employment growth by age groups. Clearly, younger firms tend to experience higher growth. Firms that are less than 5 years old grow 12.4% a year, while firms aged above 35 shrink 3.8% annually.[6]

Exploiting the panel nature of the ASIE, we estimate firm-level productivity, which allows us to relate entry directly to productivity growth. We regress firm's annual productivity growth on the age of firm, controlling for 2-digit industry fixed effects as well

[6]This pattern of declining growth with age also holds when we use firm's revenue, value added or productivity instead of employment.

Figure 2.1: Predicted Annual Firm Employment Growth by Age Groups



Note: This figures shows the predicted annual firm employment growth by age groups. Firm's employment growth is calculated as the average of annual growth rates over 2005-7. We regress firm-level employment growth rate on firm's age, controlling for employment, capital, export status, and 4-digit industry fixed effects with standard errors clustered at 2-digit industry level to obtain the predicted values.

as firm employment, capital and export status (Column [1]-[2] of Table 2.3). Consistent with the results on employment growth above, firm's productivity growth also slows down as it ages. On average, firm's aging by one year lowers the productivity growth rate by 0.3 percentage points. After merging the 4-digit industry-level entry rate from the 2004 Census with the firm-level productivity growth from the 2004-7 ASIE sample, we are ready to show how entry affects productivity growth in Column [3]-[4] in Table 2.3. More specifically, we regress firm-level annual productivity growth from 2005 to 2007 on 4-digit industry entry rate in 2004, controlling for 2-digit industry fixed effect and firm's characteristics (employment, capital, and export status). One percentage point increase in entry in 2004 increases significantly the firm-level annual productivity growth by 0.4 percentage points in subsequent years.

Table 2.3: Industry Entry, Firm Age and Firm Annual Productivity Growth

	(1)	(2)	(3)	(4)
average entry rate			0.468***	0.444**
			(6.32)	(2.09)
firm age	-0.00323***	-0.00355***	-0.000489	-0.00141*
	(-21.18)	(-7.60)	(-1.11)	(-2.01)
average entry rate × firm age			-0.0326***	-0.0251***
			(-6.58)	(-3.93)
R^2	0.007	0.011	0.007	0.011
2-digit industry F.E.	No	Yes	No	Yes
2-digit industry clustered S.E.	No	Yes	No	Yes
Observations	314032	314032	314032	314032

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: This table shows the results of regressing firm's productivity growth on firm age (column [1] and [2]) and on industry-level entry rate, firm age and their interactions (column [3]-[4]). We also include firm employment, real capital and export status in all the specifications. Entry rate is defined as the ratio of the average number of firms established in 2003 and 2004 to the total number of incumbent firms in 2004, weighted by employment. Firm productivity growth is calculated annually for the period 2005-2007. Appendix A.2 provides details on the estimation of productivity.

3 Model

The representative household has the following preference[7]

$$U = \int_0^\infty e^{-\rho t}[\ln Y(t) - L(t)]dt,$$

where $Y(t)$ is an aggregate consumption index defined as

$$\ln Y(t) = \int_0^\zeta \ln y_\nu(t)d\nu + \int_\zeta^1 \ln y_\nu(t)d\nu, \qquad \zeta \in (0,1)$$

where $y_\nu(t)$ is the output of industry $\nu \in [0,1]$. Industries are divided into two sectors: sector 1 for $\nu \in [0,\zeta]$, and sector 2 for $\nu \in [\zeta,1]$. We use $s, s = 1,2$ to denote a sector. The two sectors differ in the entry cost, which we will specify below, and are the same in all other dimensions. Each industry consists of two firms. The final industry output is an

[7]As in Aghion et al. (2001) we use a log-linear utility function to eliminate equilibrium effect on innovation through wage, and focus on the competition effect.

aggregation over outputs of the two firms,

$$y_\nu(t) = [y_{\nu,1}(t)^\delta + y_{\nu,2}(t)^\delta]^{1/\delta}.$$

The elasticity of substitution between outputs of the two firms in the same industry is governed by the parameter δ.

Use labor as numeraire, and normalize wage as 1. Under the utility function, we have that total expenditure PY always equals 1.[8] As a result, the households optimally spend 1 on each of the intermediate good. Furthermore, we can derive the demand functions of the two firms in any industry, which are

$$y_1 = \frac{p_1^{1/(\delta-1)}}{p_1^{\delta/(\delta-1)} + p_2^{\delta/(\delta-1)}}, \qquad y_2 = \frac{p_2^{1/(\delta-1)}}{p_1^{\delta/(\delta-1)} + p_2^{\delta/(\delta-1)}}.$$

Firms use labor as the only input in production. There is a quality ladder. Denote n_1 and n_2 as the positions of firm 1 and firm 2 on the ladder and denote λ as the step size. Accordingly, their productivity levels are given by $z_1 = \lambda^{n_1}$ and $z_2 = \lambda^{n_2}$. It follows that $c_1 = \lambda^{-n_1}$ and $c_2 = \lambda^{-n_2}$ are the marginal costs for firm 1 and firm 2, respectively.

The two firms in an industry engage in Bertrand competition.[9] Given the demand functions above, the optimal pricing rule follows $p_i = \frac{\epsilon_i}{\epsilon_i - 1} c_i$, where ϵ_i is the price elasticity of demand for firm $i = 1, 2$. It can be easily shown that this elasticity takes the form $\epsilon_i \equiv \frac{1-\delta\omega_i}{1-\delta}$, with $\omega_i \equiv p_i y_i = \frac{p_i^{\delta/(\delta-1)}}{p_1^{\delta/(\delta-1)} + p_2^{\delta/(\delta-1)}}$ being the revenue of firm $i = 1, 2$. Correspondingly, the profit of firm i is $\pi_i = \frac{\omega_i}{\epsilon_i}$, for $i = 1, 2$. Note that as the revenues, ω_i, are only determined by the price ratio, p_1/p_2, so are the elasticity of demand, ϵ_i. From the optimal pricing rule, it follows that the price ratio, p_1/p_2, is entirely determined by the relative cost ratio, c_1/c_2, and ultimately it is the cost ratio that matters for the price ratio, the revenues, the elasticity of demand, and the profits.

Figure 3.1 presents firm revenue as a function of the quality gap.[10] The function follows a logistic distribution, that is, it is convex initially and turns to concave eventually. The

[8] Note the Hamiltonian is $H = lnY - L + \lambda[rA + L - PY]$. From the two first order conditions, $1 = \lambda$ and $1/Y = \lambda P$, it follows that $PY = 1$.

[9] We can alternatively assume Cournot competition, under which firm $i's$ optimal pricing rule is $p_i = \frac{1}{\delta(1-\omega_i)} c_i$. The key property that revenue and profit are logistic functions in technology gaps is unchanged.

[10] Profit, which determines firm's innovating incentives, follows a similar distribution.

incremental revenue for a follower in an industry with a large gap is small; it increases as the follower catches up and it peaks when it is on par with the leader, and eventually decreases as it becomes the new leader and its quality advantage expands. To the extent that the incremental revenue affects firms' innovate efforts, firms in industries with a smaller gap, i.e. more competitive industries, have a larger incentive to innovate to escape competition.

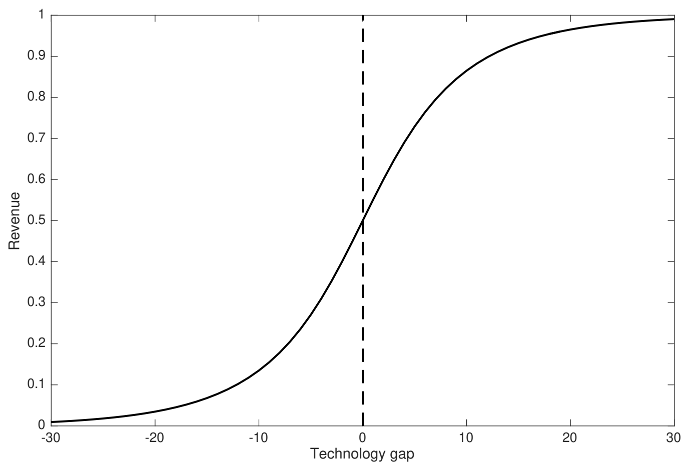


Figure 3.1: Revenue Function, Model

Note: This figure shows the revenue of a firm as a function of its quality gap relative to its opponent in the model.

Innovation In each industry, there exists a leader, a follower, and a potential entrant. Denote n the quality gap between the current leader and follower in an industry, and $\pi(n)$ and $\bar{\pi}(n)$ the associated profit for the leader and follower, respectively. We label an industry where $n = 0$ a neck-and-neck industry. When a leader innovates and succeeds, it enlarges its advantage from n to $n + 1$. Upon a successful innovation of a follower, with probability ϕ, it immediately catches up with the leader and closes completely the quality gap, i.e. from n to 0; with probability $1 - \phi$, it cuts the quality gap by 1 step, from n to $n - 1$. If a potential entrant succeeds, it replaces the follower in an industry with positive gap, i.e. $n \geq 1$, and replaces each incumbent firm with equal probability in a neck-and-neck industry.[11]

Firms are heterogeneous and have two types: High and Low. High (low) type firms have low (high) innovation cost summarized by the parameter β_i and we have $\beta_h < \beta_l$. Firms

[11]This setting for entrants is the same as Akcigit and Ates (2019).

start with high type upon entry. Overtime, a high-type firm may transit to become a low type at Poisson rate σ, while a low type is an absorbing state. This captures the fact that some firms become less productive as they grow old over time.[12,13] Depending on the types of the leader and follower pair, we can divide industries into four categories: hh, hl, lh, and ll. The first letter stands for the leader's type and the second the follower's type. For example, in an hh industry, both leader and follower are of high types. An industry is fully characterized by (i, j, n), where i and j are the types of the leader and the follower, respectively, and n is the quality gap. Use X and $\overline{X}$ to differentiate objects for the leader and the follower. In a neck-and-neck industry $(i, j, 0)$, we use X^i and X^j to differentiate from the two incumbent firms. Given our assumption of the type transition, the Poisson rate of type transition for a type-i firm is

$$\sigma_i = \begin{cases} \sigma, & \text{if } i = h; \\ 0, & \text{if } i = l. \end{cases}$$

There are two sets of value functions. The first set describes the values of the leader, the follower and the potential entrant in an industry with a quality gap of $n \geq 1$: $V_{ij}(n)$, $\overline{V}_{ij}(n)$, and $V^e_{ij}(n)$. The second set of value functions describes the values of the two incumbents and the potential entrant in a neck-and-neck industry: $V^h_{hi}(0)$, $V^l_{li}(0)$, and $V^e_{ij}(0)$. Since there is no notion of leader or follower in the neck-and-neck state, the bar notation no longer applies and the order of the types in the subscripts has no meaning. Instead, $V^h_{hi}(0)$ (or $V^l_{li}(0)$), for $i = h, l$, simply denotes the high-type (or low-type) incumbent in the neck-and-neck industry with composition $\{h, i\}$ (or $\{l, i\}$). Since $V^h_{hl}(0) = V^h_{lh}(0)$ and $V^l_{hl}(0) = V^l_{lh}(0)$, for convenience we use $V^h_{hl}(0)$ and $V^l_{lh}(0)$ in these cases. We outline the value functions as follows.

Start with the first set of value functions for an industry characterized by (i, j, n), where

[12] Acemoglu et al. (2018) make a similar assumption. We can alternatively assume that entrants draw a probability between high and low types. The important assumption here is that high type might transit to low type which is an absorbing state.

[13] The model abstract away from capital and labor market frictions, which can impact the rate at which a firm grows. In other words, the effects of such frictions are captured in a reduced form by parameters such as the quality step size, λ, and the cost of quality improvement, β_i, $i = h, l$.

$i, j \in \{h, l\}$ and $n \geq 1$. The value function for the leader is,[14]

$$
\begin{aligned}
rV_{ij}(n) = \max_{x_{ij}(n)} \; & \underbrace{\pi(n)}_{\text{profit}} - \underbrace{\beta_i \frac{x_{ij}(n)^\alpha}{\alpha}}_{\text{R\&D cost}} + \underbrace{x_{ij}(n)[V_{ij}(n+1) - V_{ij}(n)]}_{\text{successful innovation}} + \underbrace{\sigma_i[V_{lj}(n) - V_{ij}(n)]}_{\text{change of self-type}} \\
& + \underbrace{\sigma_j[V_{il}(n) - V_{ij}(n)]}_{\text{change of follower type}} + \underbrace{\overline{x}_{ij}(n)\{\phi[V^i_{ij}(0) - V_{ij}(n)] + (1-\phi)[V_{ij}(n-1) - V_{ij}(n)]\}}_{\text{successful innovation by follower}} \\
& + \underbrace{x^e_{ij}(n)\{\phi[V^i_{ih}(0) - V_{ij}(n)] + (1-\phi)[V_{ih}(n-1) - V_{ij}(n)]\}}_{\text{successful innovation by entrant}}
\end{aligned}
$$

The leader optimally chooses its innovation intensity, $x_{ij}(t)$, at the associated cost $\beta_i \frac{x_{ij}(n)^\alpha}{\alpha}$. The flow value of a leader consists of: static profit minus innovation cost; gains in value upon a successful innovation; changes in value due to an exogenous change of own type or that of the follower; and changes in value due to successful innovation by the follower or entrant.

The value function for the follower in industry (i, j, n) is

$$
\begin{aligned}
r\overline{V}_{ij}(n) = \max_{\overline{x}_{ij}(n)} \overline{\pi}(n) - \beta_j \frac{\overline{x}_{ij}(n)^\alpha}{\alpha} & + \overline{x}_{ij}(n)\{\phi[V^j_{ji}(0) - \overline{V}_{ij}(n)] + (1-\phi)[\overline{V}_{ij}(n-1) - \overline{V}_{ij}(n)]\} \\
& + \sigma_i[\overline{V}_{lj}(n) - \overline{V}_{ij}(n)] + \sigma_j[\overline{V}_{il}(n) - \overline{V}_{ij}(n)] + x_{ij}(n)[\overline{V}_{ij}(n+1) - \overline{V}_{ij}(n)] \\
& + x^e_{ij}(n)[0 - \overline{V}_{ij}(n)].
\end{aligned}
$$

Symmetrically, the flow value of a follower consists of: static profit minus innovation cost; gains in value upon a successful innovation; changes in value due to an exogenous change of own type or that of the leader; and changes in value due to successful innovation by the leader or entrant.

The value of the potential entrant in industry (i, j, n) is

$$
V^e_{ij}(n) = \max_{x^e_{ij}(n)} \; -\tau_s \beta_h \frac{x^e_{ij}(n)^\alpha}{\alpha} + x^e_{ij}(n)[\phi V^h_{hi}(0) + (1-\phi)\overline{V}_{ih}(n-1)].
$$

A successful entrant replaces the follower and catches up with the leader. The parameter $\tau_s, s = 1, 2$, stands for the entry cost. A larger τ implies less entry, which represents

[14]The sector $s, s = 1, 2$ is also a state variable for firm's value function. To simplify notation, we drop this dependence whenever it causes no confusion.

a higher entry barrier. Without loss of generality, we assume $\tau_1 > \tau_2$, that is, sector 1 features a higher barrier to entry. Note that though we do not write explicitly the dependence of the value function on $s, s = 1, 2$, the value of entrant depends on the sector it belongs to, as the entry cost is sector specific.

Similarly, we can write down the second set of value functions for firms in a neck-and-neck industry. In a neck-and-neck industry, the two incumbents obtain the same profit, denoted by $\pi(0)$. For an incumbent firm of type i, the value function is

$$\begin{aligned} rV_{ij}^i(0) = \max_{x_{ij}^i(0)} \pi(0) - \beta_i \frac{x_{ij}^i(0)^\alpha}{\alpha} + x_{ij}^i(0)[V_{ij}(1) - V_{ij}^i(0)] + \sigma_i[V_{lj}^l(0) - V_{ij}^i(0)] \\ + \sigma_j[V_{i\ell}^i(0) - V_{ij}^i(0)] + x_{ji}^j(0)[\overline{V}_{ji}(1) - V_{ij}^i(0)] \\ + x_{ij}^e(0)\left\{\frac{1}{2}\left[0 - V_{ij}^i(0)\right] + \frac{1}{2}\left[\overline{V}_{hi}(1) - V_{ij}^i(0)\right]\right\}. \end{aligned}$$

The value function for firm j can be expressed in a symmetric way.

In a neck-and-neck industry, when an entrant successfully enters, it replaces either of the two incumbents with equal probability and becomes a leader with one step ahead of the opponent. For an entrant in a neck-and-neck industry, the value is[15]

$$V_{ij}^e(0) = \max_{x_{ij}^e(0)} -\tau_s \beta_h \frac{x_{ij}^e(0)^\alpha}{\alpha} + x_{ij}^e(0)\left[\frac{1}{2}V_{hi}(1) + \frac{1}{2}V_{hj}(1)\right].$$

Stationary Distribution We focus on the balanced growth path (BGP) of the model economy. In the BGP, the distribution over industry types is stationary, Denote $\mu_{ij}^s(n)$ the fraction of industries of $(i, j, n), i, j \in \{h, l\}$ and $n \geq 0$ in sector $s, s = 1, 2$, in stationary distribution. Naturally

$$\sum_s \sum_i \sum_j \sum_n \mu_{ij}^s(n) = 1.$$

As the entry cost is fixed across industries as well as over time within each sector $s, s = 1, 2$, we can derive the stationary distribution within each sector first and then obtain the economy-wide distribution. The distribution within each sector should satisfy

$$\sum_i \sum_j \sum_n \mu_{ij}^1(n) = \zeta, \qquad \sum_i \sum_j \sum_n \mu_{ij}^2(n) = 1 - \zeta.$$

[15]Again, this value depends on the sector $s, s = 1, 2$ an entrant is in.

As the *inflow* and *outflow* across industries within each sector are symmetrical, we focus here only on sector 1 without loss of generality. To obtain the economy-wide stationary distribution, we only need to go through the same process (for different entry cost τ_2), and then aggregate across sectors. Without causing confusion, we also drop the superscript s in μ to save notation in the analysis below.

Table 3.1 lists the inflow into and outflow from an industry where both the leader and the follower are of high type, i.e. $(i = h, j = h)$, as a function of the quality gap, n. For $n = 0$, the inflow is contributed by h-type firms which were previously a follower or an entrant in an industry with gap n and successfully caught up with the then high-type leader. For $n = 1$, the inflow is contributed by high-type firms which were previously an incumbent or an entrant in a neck-and-neck industry and successfully innovated. For $n \geq 2$, the inflow comes from previously high-type leaders in the industry with gap $n - 1$ who successfully innovated. On the other hand, for all $n \geq 0$, the outflow consists of successful innovation by either incumbents or entrant, and exogenous changes in the type of the incumbents. In a stationary distribution, *inflow* is equal to *outflow* for all states. We relegate the analogous tables for industries with $(i = h, j = l), (i = l, j = h)$ and $(i = \ell, j = l)$ to Tables C.1, C.2, and C.3 in the appendix.

Table 3.1: Inflow and Outflow in Industry $(i = h, j = h)$, Sector 1

State	*Inflow*		*Outflow*
n=0:	$\sum_{n\geq 2}\left[\mu_{hh}(n)\overline{x}_{hh}(n) + \mu_{hh}(n)x^e_{hh}(n) + \mu_{hl}(n)x^e_{hl}(n)\right]\phi +$ $\mu_{hh}(1)\overline{x}_{hh}(1) + \mu_{hh}(1)x^e_{hh}(1) + \mu_{hl}(1)x^e_{hl}(1)$	$=$	$\mu_{hh}(0)\left[2x^h_{hh}(0) + x^e_{hh}(0) + 2\sigma\right]$
n= 1:	$\mu_{hh}(0)\left[2x^h_{hh}(0) + x^e_{hh}(0)\right] + \mu_{hl}(2)x^e_{hl}(2)(1-\phi) +$ $\mu_{hh}(2)[\overline{x}_{hh}(2) + x^e_{hh}(2)](1-\phi) + \mu_{hl}(0)x^e_{hl}(0)/2$	$=$	$\mu_{hh}(1)\left[x_{hh}(1) + \overline{x}_{hh}(1) + x^e_{hh}(1) + 2\sigma\right]$
n$\geq$ 2 :	$\mu_{hh}(n-1)x_{hh}(n-1) + \mu_{hl}(n+1)x^e_{hl}(n+1)(1-\phi) +$ $\mu_{hh}(n+1)[\overline{x}_{hh}(n+1) + x^e_{hh}(n+1)](1-\phi)$	$=$	$\mu_{hh}(n)\left[x_{hh}(n) + \overline{x}_{hh}(n) + x^e_{hh}(n) + 2\sigma\right]$

Note: This table lists the inflow to and outflow from all possible states (i.e. gap sizes) given a (h, h) leader-follower configuration.

Aggregate Growth The aggregate growth rate in the stationary equilibrium is the sum of growth rates in the two sectors,

$$g = g_1 + g_2.$$

As shown in Appendix C, the growth rate in sector 1 is

$$g_1 \equiv \frac{d \ln Y_1}{dt} = \left[\sum_{i=h,l} \sum_{j=h,l} \sum_{n \geq 1} \mu_{ij}(n) x_{ij}(n) + \mu(0)x(0) \right] * \ln \lambda$$

where

$$\mu(0)x(0) \equiv \sum_{i=h,l} \mu_{ii}(0) \left(2x_{ii}^{i}(0) + x_{ii}^{e}(0) \right) + \mu_{hl}(0) \left(x_{hl}^{h}(0) + x_{hl}^{l}(0) + x_{hl}^{e}(0) \right)$$

is the share of neck-and-neck industries times firm's innovation intensities in these industries. Here again μ and x refer to mass and innovation intensity in sector 1, and we omit the sector superscript for simplicity. The aggregate growth rate is equal to the average of leader's productivity growth rates for all industries with positive gap, plus average productivity growth rates for all firms in neck-and-neck industries.[16] The growth rate in sector 2, g_2, is determined in a similar way. The average growth rate in a single industry in sector 1 is then $\frac{g_1}{\zeta}$, and $\frac{g_2}{1-\zeta}$ for an industry in sector 2.

4 Quantitative Analysis

To numerically solve the model, we set a limit to the number of steps a leader can possibly be ahead of its follower and denote it by $\overline{n}$. At $n = \overline{n}$, a leading firm simply stops innovation. We verify that firms' innovation intensity in an industry with gap $\overline{n} - 1$ is indeed very close to 0.

4.1 Calibration

We calibrate the stationary equilibrium of the model to data moments in 2004-7. There are 11 parameters: $\{\rho, \zeta, \alpha, \beta_h, \beta_l, \tau_1, \tau_2, \sigma, \delta, \phi, \lambda\}$. We set $\rho = 0.03$ to match an annual interest rate of 3%. For the parameter α, which is the inverse of the cost elasticity of innovation,

[16] This is a property for this class of models (e.g. Aghion et al. (2001); Liu et al. (2019)). We refer interested readers to Appendix C for the derivation of the growth rate formula in our model.

we choose $\alpha = 2$. This is consistent with a cost elasticity of innovation of 0.5 estimated in the micro-econometric innovation literature (e.g. Blundell et al. (2002)) and that adopted in the Schumpeterian growth literature (e.g. Acemoglu et al. (2018)).

The remaining nine parameters are chosen to match model moments with those in data. The model has two sectors: Sector 1 for industries in $[0, \zeta)$ and Sector 2 for $[\zeta, 1]$, with the former featuring a larger entry barrier. We define the empirical counterpart of industries in Sector 1 (or 2) as the 4-digit industries whose entry rates are below (or above) the median entry rate of their corresponding 2-digit industry. Under the assumption of common technology within a 2-digit industry, then the dispersion of entry rates within a 2-digit industry reflects the heterogeneity in entry barriers.

Table 4.1: Parameter Values

Parameter	Description	Value
externally calibrated		
ρ	discount rate	0.03
α	inverse of innovation elasticity	2
internally calibrated		
ζ	size of Sector 1	0.5
β_h	innovation cost of high type firms	0.64
β_l	innovation cost of low type firms	1.63
τ_1	entry cost in Sector 1	2.43
τ_2	entry cost in Sector 2	1.11
σ	high-to-low type transition rate	0.19
δ	elasticity of substitution within industry	0.73
ϕ	probability of drastic innovation	0.11
λ	quality step	1.23

Note: This table lists the externally calibrated parameter values and the internally calibrated parameter values.

The parameters β_h and β_l reflect the cost of innovation for the high type and low type firms, respectively. In the model, leaders and followers of different types have different costs of and returns to innovation, and therefore choose quite different innovation intensities. Due to the unobserved types, these dimensions cannot be directly mapped to data. Instead we select moments based on observable firm characteristics such as firm size and age to discipline the model. We label firms in the ASIE 2004-7 panel whose revenue is above (below) the industry median as large (small) firms and firms whose age is above (below) the industry median as old (young) firms. Following Acemoglu et al. (2018), we simulate $10,000$ industries for $3,500$ periods and calibrate β_h and β_l such that the average output growth rates of young and old leaders in the model are consistent with the average annual output growth rates of young and old firms in data from 2004 to 2007. Intuitively, due to the assumption on the type transition process, old leaders tend to have a low type while young leaders are of a high type with high probability, therefore the growth margin between the two is informative about the type-specific innovation costs. The two entry cost parameters τ_1 and τ_2 are chosen to match the entry rates in model with their empirical counterparts in both sectors.[17]

The parameter σ, which governs the transition rate from high to low type, is chosen such that the model simulated probability of transiting from large to small firms within one year matches its counterpart in the ASIE 2004-7 panel. The elasticity of substitution between firms within the same industry, δ, determines the division of total revenue between profit and wage, and is chosen to match the average labor share.[18] The probability of drastic innovation, ϕ, directly affects the size of entrants and their initial growth. We choose the value of ϕ to match the model simulated probability of remaining small for entrants after 1 year with its ASIE counterpart. Last, the quality step parameter λ is set to match the average annual output growth rate from 2004-7. These parameters capture, in a reduced-form way, how the existing institutions in 2004-7 (e.g. capital and labor market institutions and the legal environment) support economic growth.

We have a total of nine moments to calibrate the nine parameters internally. After computing the model moments from simulated data, we choose parameter values to minimize

[17]Note that in the model the high entry rate and low entry rate sectors differ only in entry barriers. We do not assume ex ante heterogeneity (in terms of $\beta's$ in the model) between firms in the two sectors. Therefore the observed difference in e.g. firm growth and market power between the two sectors is only an endogenous response to different degree of entry regulation and competition. Put it differently, if one pulls a firm from Sector 1 (which we will provide an empirical interpretation of a SOE-dominated sector) and plug it in Sector 2, it will behave in the same way as those firms which have originally entered Sector 2.

[18]The calibrated value of δ, 0.73, implies a within-industry elasticity of substitution between leader and follower of $\frac{1}{1-\delta} = 3.70$. We have tried larger values e.g. $\delta = 0.85$ and 0.9. The effect of δ, while keeping all other parameters unchanged, on the aggregate growth rate in our model is nonlinear. A larger value of δ increases leader's profits and innovation intensity, but also dampens entry. However, the sectorial differences presented below is robust to different values of δ.

Table 4.2: Data and Model Moments

Moment	Data	Model
size of sector 1	0.500	0.500
growth rate of young firms	0.139	0.109
growth rate of old firms	0.066	0.047
1 year entry rate in sector 1	0.090	0.090
1 year entry rate in sector 2	0.120	0.143
large-to-small transition probability	0.066	0.044
unweighted mean of LS	0.500	0.502
probability of small for entrants	0.625	0.699
aggregate growth rate	0.090	0.090

Note: This table lists the targeted moments in the data and their counterparts produced by the calibrated model.

a weighted sum of the distance between model and data moments:

$$\sum_{k=1}^{9} \iota_k \frac{|\text{model}(k) - \text{data}(k)|}{0.5 * |\text{model}(k)| + 0.5 * |\text{data}(k)|}$$

To match well at the macro level, the moments of aggregate output growth rate is assigned a weight (ι_k) 5 times the weight of others. Table 4.1 summarizes the calibrated parameter values and Table 4.2 lists the moments used in the calibration.

4.2 Sector Heterogeneity

Under the calibrated parameters, the average growth rate of output in Sector 1 is 8.02% and the average growth rate in Sector 2 is 9.98%. Note that the only difference between Sector 1 and Sector 2 in the model is in the entry cost. A lower entry cost affects growth along four margins. One, it induces more innovation efforts from potential entrants and results in industries with on average younger firms who tend to growth faster, i.e. a direct positive effect on growth. Two, it discourages incumbents from costly innovation in a given market structure under the threat of more entry, i.e. a negative Schumpeterian effect on growth. Three, it improves the endogenous distribution of firms' types, since high-type entrants replace potentially low-type incumbents. Four, it changes the endogenous distribution of firms over different industry-level market structures, essentially relocating firms towards industries which are more competitive with closer quality gaps between the leader and follower. This last effect is a growth-enhancing effect and, as our growth decomposition exercise shows, turns out to be the most important channel through which entry promotes growth.

One way to visualize how the different channels are at work, which benefits from having direct data counterpart, is to examine how age and measures of competitiveness are distributed across industries in the two sectors. As in the empirical section, we use HHI to measure the competitiveness of an industry. Recall that ω_1 and ω_2 are both the revenue and revenue share of the two firms in an industry. As a result, the industry HHI is

$$HHI = \omega_1^2 + \omega_2^2.$$

As both ω_1 and ω_2 are functions of the quality gap n, the industry HHI also depends on n. In particular, it is straightforward to show that a larger n corresponds to a larger HHI. We simulate a large sample of industries and firms for long enough to reach the BGP with a stationary distribution of firms' age and industry attributes and plot the distribution of the average firms' age and HHI over industries across the two sectors in Figure 4.1.

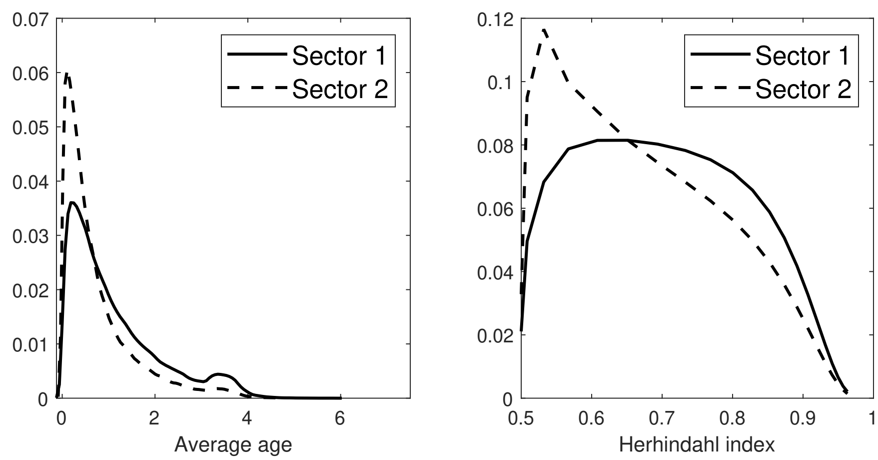
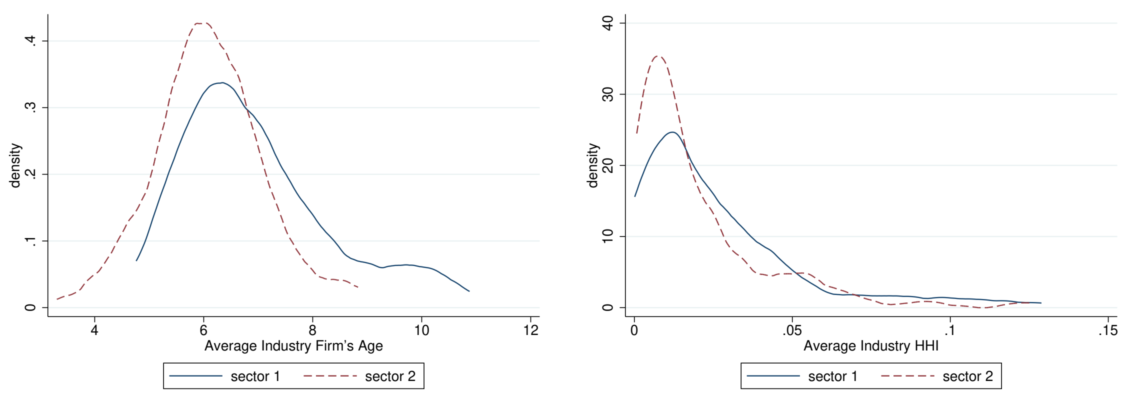


Figure 4.1: Average Age and HHI Distribution in Sector 1 and 2, Model

Note: This figure shows the distribution of industry-level average firm age in Sector 1 and 2 respectively (Panel (a)) and the distribution of industry-level HHI in Sector 1 and 2 respectively (Panel (b)) from the model simulated data.

Compared to Sector 1 which has a larger entry cost, there are more young firms in Sector 2 which has more entry. In the model, young entrants replace old incumbents upon entry, a higher entry rate necessarily leads to an age distribution that is more skewed to the left, as show in the left panel of Figure 4.1. On the other hand, compared to Sector 2, Sector 1 with less entry end up having more industries in which the leading firm becomes dominant and faces little challenge from either the follower or an entrant. In those industries, the concentration is high, and since there is little incentive to innovate for both the leader and follower, the growth rate is low. We observe a thicker right tail in the distribution of industry concentration proxied by HHI in Sector 1 in the right panel of the same figure.

Even though we do not directly target these endogenous distributions to match those in the data, the data counterparts portray a similar picture. Recall that we classify all 4-digit manufacturing industries in 2004 Industrial Census whose entry rate is below (above) the median entry rate in the corresponding 2-digit industry as the data counterpart of industries belonging to Sector 1 (2). Figure 4.2 shows the distribution of average age and the HHI across industries in Sector 1 and Sector 2 in the data. Clearly, Sector 2 have more industries with younger firms, while Sector 1 have more industries with high concentration. It's worth pointing out that because our model features only two firms per industry, the model necessarily misses the absolute levels of age and HHI in the data. However, the model does a reasonably good job in terms of replicating the qualitative differences across the two sectors caused by differential entry costs.



(a) Firm Age

(b) HHI

Figure 4.2: Distribution of Firm Age and Industry HHI by Sectors, 2004 Census Sample

Note: This figure shows the empirical distribution of industry-level average firm age in the high and low entry sector respectively (Panel (a)) and the empirical distribution of industry-level HHI in the high and low entry sector respectively (Panel (b)) from the 2004 Industrial Census sample.

Growth Decomposition We present a decomposition of the gain of growth in Sector 2 relative to that in Sector 1 to organize our thoughts around the underlying mechanisms at work. We have derived in Section 3 the formula for the aggregate growth rate of the economy. To conserve notation, use ψ to denote the type configuration of a leader-follower pair, i.e. $\psi = (h,h), (h,l), (l,h), (h,h)$; and keep $n \in \{0,1,...,\overline{n}\}$ to denote industry's quality gap. We rewrite the growth rate formula for sector s as

$$g_s = \sum_{\psi} \mu_s(\psi,0) x_s^e(\psi,0) \ln \lambda + \sum_{\psi} \sum_{n} \mu_s(\psi,n) x_s(\psi,n) \ln \lambda, \quad s = 1,2.$$

where the first component denotes all terms associated with the entrant's innovation intensity in the growth rate formula, and the second component denotes all remaining terms in the growth rate formula.

In the model, Sector 2 has a lower entry barrier and a higher growth rate. We can decompose the effect of a lower entry cost on the gain in aggregate growth in Sector 2 as follows[19]

$$\begin{aligned} g_2 - g_1 \approx & \underbrace{\sum_{\psi} \mu_1(\psi,0)[x_2^e(\psi,0) - x_1^e(\psi,0)] \ln \lambda}_{\textit{direct effect}} + \underbrace{\sum_{\psi} \sum_{n} \mu_1(\psi,n)[x_2(\psi,n) - x_1(\psi,n)] \ln \lambda}_{\textit{Schumpeterian effect}} \\ & + \underbrace{\sum_{\psi} \sum_{n} x_2(\psi,n)[f_2(\psi|n) - f_1(\psi|n)] \tilde{\mu}_1(n) \ln \lambda}_{\textit{replacement effect}} \\ & + \underbrace{\sum_{\psi} \sum_{n} x_2(\psi,n)[\tilde{\mu}_2(n) - \tilde{\mu}_1(n)] f_2(\psi|n) \ln \lambda}_{\textit{pro-competitive effect}} . \end{aligned}$$

where $f_s(\psi|n), s = 1,2$ denotes the distribution of ψ conditional on a given value of n, and $\tilde{\mu}_s(n) \equiv \sum_{\psi} \mu_s(\psi,n)$ is the (unnormalized) marginal distribution of n in sector $s, s = 1,2$.

We can decompose the effect of higher entry on aggregate output growth into four components: a direct effect, a Schumpeterian effect, a replacement effect and a pro-competitive effect. As entrants' innovation intensity in neck-and-neck industries directly enters the growth rate formula, a lower entry cost directly increases this intensity and therefore promotes aggregate growth. When the entry rate is high, incumbent leaders are more likely to face a high type challenger and consequently face a higher probability of being overtaken, which discourages the incumbent leader to invest in innovation. This is the Schumpeterian effect, typical in models of creative destruction, and it dampens growth.[20] Both

[19] As detailed in Appendix C, there are two symmetric approaches to decompose the changes in $\mu's$ into the replacement effect and the pro-competitive effect. The relative magnitude of the two effects differ non-trivially under the two approaches. In the text, we report the average from these two approaches.

[20] A closer look at the difference, $x_2(n) - x_1(n)$, at various levels of n reveals that there is also a secondary

the direct effect and the Schumpererian effect work through the x terms in the growth rate formula.

A higher entry rate also changes the industry composition, i.e. the μ terms. Firms become less productive stochastically over their lifetime. When there are more entrants, more young and productive firms enter and replace old and (on average) less productive incumbents. In other words, the distribution of the four type configurations of leader-follower pairs, ψ, given any quality gap n, or $f(\psi|n)$, evolves in such a way that relocates industries away from (l, l) and towards (h, h). This is the replacement effect, which tends to increase aggregate growth. Lastly, when the entry barrier is lower, it is more difficult for an incumbent firm to accumulate and build up advantage. The economy thus have more industries in which the quality distance between firms are close and competition is fierce. Since both firms in more competitive industries innovate more and a lower entry barrier shifts more masses to such competitive industries, we refer to this last effect the pro-competitive effect and it is growth-enhancing.

The result of the decomposition is found in Table 4.3.[21] Under the calibrated parameters, the negative Schumpeterian effect almost exactly cancels out the positive direct effect associated with higher entrants' innovation intensity in neck-and-neck industries. This means the gain in growth in Sector 2 comes entirely from the compositional change of industry distribution over types and quality gaps (or competitiveness), namely the replacement effect and the pro-competitive effect. More specifically, of the 1.9 percentage point difference in the growth rates between Sector 1 and 2, about 40% is due to the replacement effect and 60% is due to the pro-competitive effect.

The replacement effect can be discerned from Table 4.4, where we show the distribution of industries over the four type configurations of leader-follower pairs for the two sectors. In Sector 2, 27.0% of industries have a high-type leader and 50.4% of industries have

effect, whereby for an incumbent leader who is having a intermediate value of lead n over the follower, its innovation effort can be higher in Sector 2 than in Sector 1. This happens because the Schumpeterian effect is especially strong when n is small and diminishes as n gets larger, so that, faced with the threat of higher entry, the incumbent leader innovates more to escape future competition. Quantitatively, this effect is so small that when summing over different n, the Schumpeterian effect clearly dominates.

[21] The sum of the four effects is slightly larger than the overall growth rate, mainly because the decomposition is only an approximation. It should be pointed out that this decomposition results are not driven by weights chosen in the decomposition formula. Using $\frac{\mu_1(n)+\mu_2(n)}{2}$ instead of $\mu_1(n)$ as the weights in the Schumpeterian effect term, and $\frac{x_1(n)+x_2(n)}{2}$ instead of $x_1(n)$ as the weights in the composition effect term, we obtained effects that are quite similar to those in Table 4.3.

Table 4.3: Decomposition of Growth Rate Differences, Model

	growth rate	direct	Schumpeterian	replacement	pro-competitive
S2-S1	0.019	0.002	-0.002	0.0087	0.0126
%		9.39%	-9.39%	40.85%	59.15%

Note: This table shows the decomposition of the growth difference between Section 1 and 2 into the direct, Schumpeterian, replacement, and pro-competitive effects.

a high-type follower; these two numbers are lower at 18.4% and 38.6% in Sector 1. The higher prevalence of the high type in Sector 2 means that Sector 2 has higher shares of industries with (h,h), (h,l) and (l,h) configurations and a lower share of industries with (l,l) compared to Sector 1. To the extent that firm's age is closely linked to its type, the left panel of Figure 4.1, which shows the distribution of age over industries by sector, confirms this patterns as well.

Table 4.4: Distribution of Industry Types, Model

	(h,h)	(h,l)	(l,h)	(l,l)	h Leader	l Follower
Sector 1	0.036	0.056	0.157	0.250	18.4%	38.6%
Sector 2	0.070	0.065	0.182	0.183	27.0%	50.4%

Note: This table shows the composition of the four types of leader-follower configurations in Sector 1 and 2 in the model.

The pro-competitive effect, which shifts mass of industries from large quality gaps towards small quality gaps, is demonstrated in the right panel of Figure 4.1. The HHI is simply an increasing transformation of quality gap n. The fact that the pro-competitive effect is growth-enhancing is shaped by the escape competition force in models with step-by-step innovation (e.g. Aghion et al. (2001)). The escape competition force is at its strongest when the distance between the leader and the follower is close. In such an industry, the follower has strong incentive to try to leapfrog the leader and as a consequence the leader has strong incentive to innovate to escape from competition. This negative relationship between leader's incentive to innovate and its advantage is depicted in Figure 4.3 for the case of (h,h) industries in Sector 2.[22] Since most of the innovation in an economy happens in relatively competitive industries, it then naturally follows that the pro-competitive ef-

[22]The initial jump corresponds to movement from $n=0$ to $n=1$. Compared to a leader that is one step ahead of its follower, incumbent firms innovate slightly less in a neck-and-neck industry. For $n \geq 1$, the leader's innovation intensity monotonically decreases in its gap n.

fect of entry boosts aggregate innovation and promotes growth.

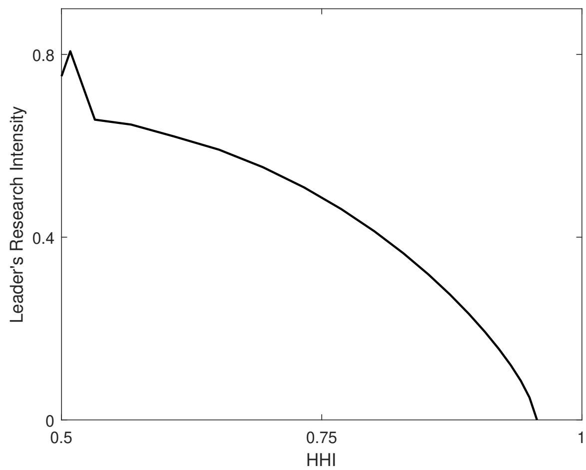


Figure 4.3: HHI and Leader's Innovation Intensity in a (h, h) Industry in Sector 2

Note: This figure shows the leader's innovation intensity as a function of HHI in Sector 2 of the model.

To sum up, in this section, we have shown how entry affects aggregate growth through the different channels in our model. In the case of China, when we compare the industries with lower cost of entry to those with higher cost of entry, the gain in output growth is entirely driven by compositional changes in how industries are distributed over types and quality gaps. In the next section, we relate the entry barriers to the SOE reforms in China and conduct a counterfactual simulation of the model to quantify the gain in aggregate growth in the entire manufacturing sector from removing entry barriers over time through a series of SOE reforms.

5 Entry Barriers and SOE Reform

Entry cost can stem from different sources. Over several decades, technological progress in transportation, information and communications, and financial technology can naturally lower the cost of entry in all sectors. This technological component of entry cost differs from what we consider as entry barriers. In the historical context of the Chinese economy, a major form of artificial entry barriers is associated with the state-imposed restriction of entry to protect the state sector. In this section, we review this interpretation of entry barrier and conduct a counterfactual analysis based on such an interpretation. The counterfactual exercise seeks to quantify the contribution to aggregate growth in the manufacturing sector from the dismantling of entry barriers by the SOE reforms which took place in the late 1990s and early 2000s.[23] In other words, we take the model calibrated to the 2004-7 Chinese manufacturing sector in the previous section and assess counterfactually what the growth rate would be if the entry barriers in 2004-7 were as high as those in 1995.

Institutional Background The Economic Reform and the Opening Up of China since the late 1970s comprise a series of economic reforms that aim to transform what was a centrally planned system with state ownership towards a market economy with diverse ownership types. Under the planned regime, the Chinese economy was dominated by state-owned enterprises, with close-to-zero entry and exit. Private firms were not allowed to enter and operate, while low-efficiency SOEs would not be pushed out. While the reform in the late 1970s and early 1980s mainly involved the de-collectivization of agriculture, the opening up to foreign investment and market reforms in a few selected areas, the second stage of SOE reform was launched in late 1980s and continued throughout 1990s, especially after Deng Xiaoping's Southern Tour in 1992.

The subsequent reform encouraged the entry of private firms, gradually across industries, and allowed non-productive SOEs to exit the market. In 1994, a new Company Law was adopted, which provides a framework for the process of converting SOEs into corporations. In 1995, the policy "grasping the large and letting go the small" (*zhuada fangxiao*) was adopted. During this process, entry barrier, competition, and market structure all experienced dramatic changes. As pointed out by Qian (2002), one important pillar of the success of the reform lies in the fact that it unleashes the standard forces of incentives and competition.

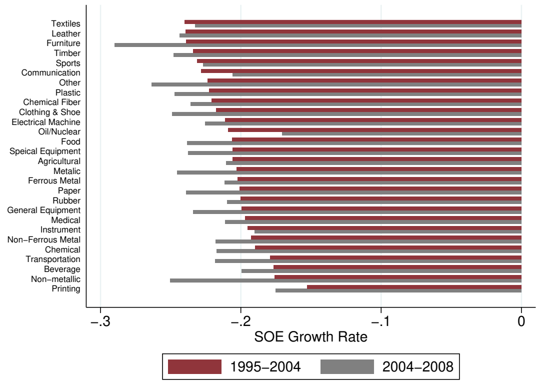
From our data source, we observe marked differences in the speed of the reduction of SOE shares across different industries, which is consistent with the logic of reform. The panel (a) of Figure 5.1 shows a clear variation in industries of changes in the SOE shares. Consistent with Li et al. (2015), the exit of SOEs was first concentrated in more downstream industries, such as manufacturing of consumption goods, and gradually spread to more upstream industries. In panel (b) of the same figure, there is clear cross-sectional

[23] During our sample period, another major change in the economic environment is China's entry to WTO. In Appendix B.3, we review evidence of the role that foreign owned enterprises and domestic exporting firms play during this time period. We show that the empirical regularities that we focus on, i.e. the increased competition and growth following a reduction in entry barrier, are unlikely driven by either foreign owned enterprises or export oriented firms.

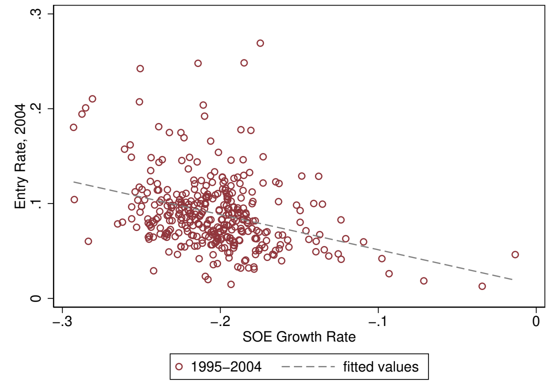
negative correlation between the growth rates of SOE shares from 1995-2004 and the entry rate in that industry in the 2004 Industrial Census. Panel (c) further shows that sectors with greater drop in SOE shares over our entire Census sample period tend to experience a faster revenue growth.[24]

[24] In Appendix B.2, Table B.2 shows that such relationship between change in SOE shares and revenue growth still holds after controlling for the initial industry characteristics such as total number of firms and average employment and 2-digit CIC industry fixed effects, which is also robust to the choice of various time periods.

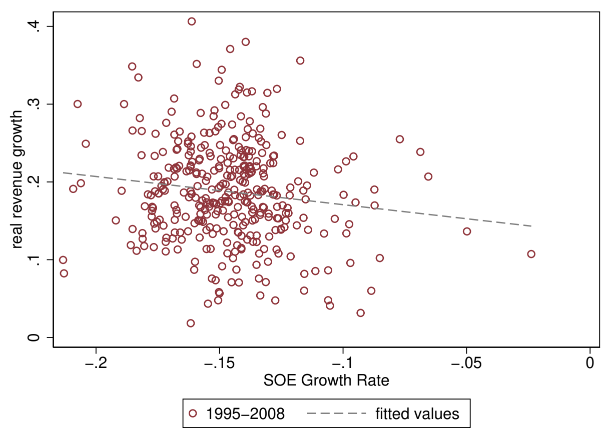
Figure 5.1: Stylized Facts on SOE Shares, Entry, and Revenue Growth



(a) SOE Growth Rate



(b) Entry and SOE Growth Rate



(c) Revenue Growth Rate and SOE Growth Rate

Note: This figure shows stylized facts related to the change in SOE shares across industries. Panel (a) shows the change in SOE share from 1995 to 2004 and from 2004 to 2008 by 2-digit CIC industry; panel (b) shows the scatter plot between 4-digit CIC industry-level SOE growth rates from 1995 to 2004 and the entry rates in 2004; panel (c) shows the scatter plots between 4-digit CIC industry-level SOE growth rates and real revenue growth rates from 1995 to 2008. Entry rate is defined as the ratio of the average number of firms established in 2003 and 2004 to the total number of incumbent firms in 2004, weighted by employment.

Counterfactual Analysis To answer the question what is the contribution to aggregate growth in the Chinese manufacturing sector from reducing entry barriers associated with the SOE reforms, we must first isolate the part of entry that is induced by policy. To that end, we make use of the elasticity of entry rates with respect to the presence of the SOEs in an industry.

Using a 3-wave panel of 4-digit industries constructed from the 1995, 2004 and 2008 Industrial Census, we regress the industry-level entry rates on the same-year industry-level share of SOEs, controlling for time-varying industry characteristics such as total revenue, employment, and number of firms as well as industry fixed effects (Column [2] of Table 5.1). We find that one percentage point increase in SOE share in an industry is associated with 0.0428 percentage points reduction in its entry rate. Since the SOE employment shares for the manufacturing sector are 85.13% in 1995 and 16.56% in 2004, the difference in the SOE presence between 1995 and 2004 would imply a 2.93 percentage points difference in entry rates between 1995 and 2004. Starting from a manufacturing sector wide of entry rate of 10.51% in 2004, the counterfactual entry rate in 1995, absent the entry-promoting effect of the SOE reforms, is 7.58%. The 1995 entry rate in data is 7.37%.

Table 5.1: Industry SOE Share and Industry Entry Rate

	(1)	(2)
soe share	-0.0245***	-0.0428***
	(-3.57)	(-5.73)
log industry revenue	-0.00418**	-0.0293***
	(-2.68)	(-5.57)
Constant	0.131***	0.391***
	(7.07)	(5.74)
R^2	0.066	0.509
4-digit industry F.E.	No	Yes
Observations	1174	1164

t statistics in parentheses
standard errors clustered at 2-digit industry level
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: This table reports the results of regressing 4-digit CIC industry-level entry rates on industry SOE shares, controlling for industry characteristics such as total number of firms and employment. Entry rate is defined as the ratio of the average number of firms 1 year before a particular year and in that particular year to the total number of incumbent firms in that particular year, weighted by employment. SOE share is the share of SOE employment in that particular year.

Collapsing two sectors into one which has a single value of τ, we recalibrate τ to target a counterfactual entry rate of 7.58% while keeping all other parameters unchanged. We reduce the two sectors to one under the interpretation that the systematic difference in entry rates across the two sectors in 2004 is primarily driven by the gradual SOE reforms across sectors that took place between 1995 and 2004. The counterfactual entry rate we obtain is interpreted as a consequence of uniformly high entry barriers across industries in the pre-reform era.

The recalibrated model results in a reduction in the aggregate growth rate from the baseline 9% to 7.57%. That is, the reduction of entry barriers achieved by the SOE reforms between 1995 and 2004 explains about 15.9% of the growth achieved during 2004-7. Using the only data source we have for earlier years, we do a back-of-the-envelope calculation using labor productivity of the industrial sector from the NBS data in Figure 1.1 and find that the increase of the average growth rate from 1991-5 to 2004-7 is about 37.8% of the level of the growth from 2004-7. Since our result suggests that the reduction of entry barriers accounts for 15.9% of the level of growth from 2004-7, this is 42% of the growth differential between 1991-5 and 2004-7. The remaining growth differential may be explained by other factors (e.g. factor market reforms) which could change parameters other than the entry cost (e.g. λ, β, ϕ or σ) from the 1991-5 economy to the 2004-7 economy.

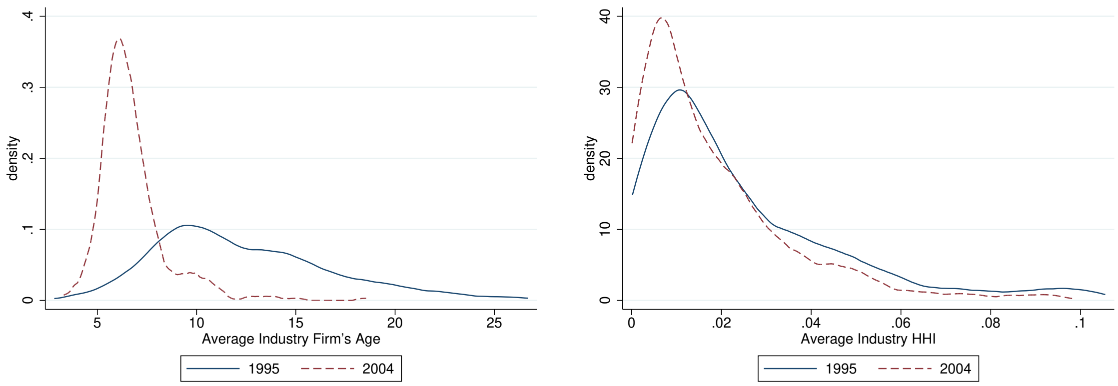
When we decompose the change in aggregate growth into the four effects as in Section 4.2, we find again the pro-competitive effect explains the majority of the gain in growth (Table 5.2). More specifically, of the 1.43 percentage points gain of aggregate growth, the direct effect explains 10.97%, the Schumpeterian effect a negative 6.45%, the replacement effect 38.71% and the pro-competitive effect 56.78%.

Table 5.2: Decomposition of the Gain in Growth Upon Entry Barriers Reduction

Δ growth rate	direct	Schumpterian	replacement	pro-competitive
0.0143	0.0017	-0.001	0.0060	0.0088
–	10.97%	-6.45%	38.71%	56.78%

Note: This table reports the result of decomposition of the growth rate difference into direct, Schumpeterian, replacement and pro-competitive effects.

The model predicts that a reduction of entry barriers leads to an economy with more young firms, consistent with the replacement effect, and more competitive industries, consistent with the pro-competitive effect, both of which have empirical counterparts. Using the 1995 and 2004 Industrial Census, we plot in Figure 5.2 the distributions of average firm age and HHI across 4-digit manufacturing industries in the 1995 and 2004 Census data. Both of the predictions are borne out by the data.



(a) Firm Age

(b) HHI

Figure 5.2: Distribution of Firm Age and Industry HHI, 1995 and 2004 Census Samples

Note: This figure shows the distribution of average firm age and HHI over 4-digit industries in the 1995 and 2004 Industrial Census respectively.

6 Conclusion

In this paper, we revisit the narrative that the gradual economic reforms removed hurdles to enter previously state-dominated industries, unleashed unprecedented competition, and achieved remarkable aggregate growth in the economic history of the People's Republic of China. We examine the empirical evidence from the three waves of Industrial Census and a ten-year panel of Chinese manufacturing firms through the lens of a model of endogenous productivity, entry and market structure with heterogeneous firms and sectors. To the best of our knowledge, we are the first to adopt such a theoretical framework to understand the effect of the reduction of entry barrier on growth in China.

We calibrate the model to the Chinese manufacturing sector in 2004-7 and use it to understand the cross-industry differences in entry rates, growth and competition in that time period. After we decompose the growth difference between the sectors with high and low entry rates, we find that higher entry contributes to higher growth mainly through compositional changes in firm types and industry competitiveness. The replacement effect, which reflects the higher prevalence of the more productive type as a result of entry, explains 40% of the growth difference. This effect resembles the selection effect in the literature which has examined the role of entry through the lens of a model of perfect competition or monopolistic competition. The pro-competitive effect, which can only be identified in our model with endogenous market structure, explains 60% of the growth difference.

When we run a counterfactual experiment and quantify the amount of growth brought by reducing entry barriers associated with the SOE reforms which occurred in late 1990s and early 2000s, we find that the entry margin alone explains more than 40% of the aggregate growth rate differential of the Chinese manufacturing sector between 1991-5 and 2004-7. A similar decomposition also identifies the pro-competition effect as the dominant channel through which entry promotes growth, explaining about 57% of the gain in growth. By focusing on entry, we inevitably abstract away from reforms in other spheres of the economy such as trade and urbanization.

While our framework permits growth to respond endogenously to changing market structure, it has limitations. For instance, it cannot generate bouts of growth of the capital deepening type that occur in the transitional dynamics of a neoclassical growth model. In other words, the framework does not feature an endogenous mechanism that can produce non-monotone growth trajectories. By calibrating the model to the 2004-7 period, we take a snapshot of the economy and use parameters such as the step size and costs of innovation to capture, in a reduced-form way, how the existing institution supports growth. Nevertheless, the counterfactual exercises are informative about the contribution of entry to the observed level of growth in 2004-7 as long as entrants are faced with the same economic environment as young incumbents. More generally, we also recognise that entry barrier is only one form of anti-competitive measures. Unequal access to credit and financial markets, preferential treatment in tax/subsidies, political interference in commercial activities or biased courts can all hinder competition and prevent the economy from achieving its growth potential. We leave each of these topics for future research.

References

Acemoglu, Daron, Ufuk Akcigit, Harun Alp, Nicholas Bloom, and William Kerr, "Innovation, Reallocation, and Growth," *American Economic Review*, 2018, *108* (11), 3450–91.

Aghion, Philippe, Christopher Harris, Peter Howitt, and John Vickers, "Competition, Imitation and Growth with Step-by-Step Innovation," *Review of Economic Studies*, 2001, *68* (3), 467–492.

_, Nick Bloom, Rchard Blundell, Rachel Griffith, and Peter Howitt, "Competition and Innovation: An Inverted-U Relationship," *Quarterly Journal of Economics*, 2005, *120* (2), 701–728.

_, Robin Burgess, Stephen Redding, and Fabrizio Zilibotti, "Entry Liberalization and Inequality in Industrial Performance," *Journal of the European Economic Association*, 2005, *3*, 291–302.

Akcigit, Ufuk and Sina T. Ates, "What Happened to U.S. Business Dynamism," *Working paper*, 2019.

Asturias, Jose, Sewon Hur, Timothy J. Kehoe, and Kim J. Ruhl, "Firm Entry and Exit and Aggregate Growth," *NBER Working Paper*, 2019.

Ates, Sina T. and Felipe E. Saffie, "Fewer but Better: Sudden Stops, Firm Entry, and Financial Selection," *American Economic Journal: Macroeconomics*, 2020. forthcoming.

Blundell, Richard, Rachel Griffith, and Frank Windmeijer, "Individual Effects and Dynamics in Count Data Models," *Journal of Econometrics*, 2002, *108* (1), 113–131.

Brandt, Loren, Gueorgui Kambourov, and Kjetil Storesletten, "Barriers to Entry and Regional Economic Growth in China," *CEPR Discussion Papers*, 2020.

_, Johannes Van Biesebroeck, and Yifan Zhang, "Creative Accounting or Creative Destruction? Firm-Level Productivity Growth in Chinese Manufacturing," *Journal of Development Economics*, 2012, *97* (2), 339–351.

_, Thomas G. Rawski, and John Sutton, "China's Industrial Development," in Loren Brandt and Thomas G. Rawski, eds., *China's Great Economic Transformation*, Cambridge University Press, 2008, p. 569–632.

Cheremukhin, Anton, Mikhail Golosov, Sergei Guriev, and Aleh Tsyvinski, "The Economy of People's Republic of China from 1953," *Working paper*, 2017.

Groves, Theodore, Yongmiao Hong, John McMillan, and Barry Naughton, "Autonomy and Incentives in Chinese State Enterprises," *Quarterly Journal of Economics*, 1994, *109* (1), 183–209.

Hall, Robert E, Olivier Jean Blanchard, and R Glenn Hubbard, "Market Structure and Macroeconomic Fluctuations," *Brookings Papers on Economic Activity*, 1986, *1986* (2), 285–338.

Haltiwanger, John, Ron S. Jarmin, Robert Kulick, and Javier Miranda, *High Growth Young Firms: Constribution to Job, Output, and Productivity Growth*, National Bureau of Economic Research, Inc., 2016.

Herrendorf, Berthold and Arilton Teixeira, "Barriers to Entry and Development," *International Economic Review*, 2011, *52* (2), 573–602.

Hopenhayn, Hugo A., "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 1992, *60* (5), 1127–1150.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

— **and Zheng Michael Song**, "Grasp the Large, Let Go of the Small: The Transformation of the State Sector in China," *Brookings Papers on Economic Activity*, 2015, pp. 295–346.

Levinsohn, James and Amil Petrin, "Estimating Production Functions Using Inputs to Control for Unobservables," *The Review of Economic Studies*, 2003, *70* (2), 317–341.

Li, Xi, Xuewen Liu, and Yong Wang, "A Model of China's State Capitalism," *Available at SSRN 2061521*, 2015.

Liu, Ernest, Atif Mian, and Amir Sufi, "Low Interest Rates, Market Power, and Productivity Growth," Technical Report, National Bureau of Economic Research 2019.

Loecker, Jan De and Frederic Warzynski, "Markups and Firm-Level Export Status," *American Economic Review*, 2012, *102* (6), 2437–71.

McMillan, John and Barry Naughton, "How to Reform a Planned Economy: Lessons from China," *Oxford Review of Economic Policy*, 1992, *8* (1), 130–143.

Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 2003, *71* (6), 1695–1725.

North, Douglass C. and Robert Paul Thomas, *The Rise of The Western World: A New Economic Histroy*, Cambridge University Press, 1973.

Parente, Stephen L. and Edward C. Prescott, "Monopoly Rights: A Barrier to Riches," *American Economic Review*, 1999, *89* (5), 1216–1233.

Peters, Michael, "Heterogeneous Markups, Growth and Endogenous Misallocation," *Econometrica*, 2020, *88* (5), 2037–2073.

Qian, Yingyi, "How Reform Worked in China," *William Davidson Institute Working Paper 473*, 2002.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, "Growing Like China," *American Economic Review*, 2011, *101* (1), 196–233.

Zhu, Xiaodong, "Understanding China's Growth: Past, Present, and Future," *Journal of Economic Perspectives*, 2012, *26*, 103–124.

FOR ONLINE PUBLICATION

Online Appendix of "Growing through Competition: The Reduction of Entry Barriers among Chinese Manufacturing Firms"

Appendix A Data and Sample Construction

A.1 ASIE Sample Construction

ASIE Panel Construction. Following Brandt et al. (2012), we create an unbalanced panel of firms between 1998 and 2007, using the unique firm IDs to link firm over time. For those that cannot be matched by IDs, we use additional information such as firm name, name of legal person representative, address, industry, and etc. Since it's possible firms exit the sample and reenter latter, according to their methodology, we first match the samples of two consecutive years, then three consecutive years, and finally create a 10-year-panel sample.

Industry Code. Firms in the ASIE and Industrial Census samples are classified into an industry by the 4-digit Chinese Industry Classification (CIC) system. CIC 1994 codes and 2002 codes are used in different years of our sample, and there was a revision of the classification system in 2003. In order to harmonise the CIC code and make it comparable over time, we follow Brandt et al. (2012) and implement the same industry concordance. For the industries that cannot be converted using their methods in Industrial Census, we manually match them using their Chinese names.

State Ownership. To define state ownership in ASIE, we follow the approach as in Hsieh and Song (2015). A firm is defined as state-owned when the share of registered capital held directly by the state is more than 50 percent of when the controlling shareholder is reported as the state. We have verified that we can replicate the main facts documented in their paper using our sample. To define state ownership in Industrial Census, we use firms' "registered ownership type." The non State-Owned Enterprises include Private-Owned Enterprises (POEs) and Foreign-Owned Enterprises (FOEs). We create a state dummy, which is 1 if a firm is a SOE and 0 if a firm is a FOE or POE.

Real Output and Input Values. Information of output and value added is missing in the year 2004 ASIE survey, and hence we have to impute output and value added for firms in 2004. We use information from other sample periods and run a simple OLS regression of log output on log revenue and obtain the predicted output for year 2004. We calculate the first-step-imputed value added as the difference between imputed output and intermediates. Then we regress log value added on log first-step-imputed value added and obtain the imputed value added for year 2004. Now we have output and value added for year sample periods. We further reconstruct the gross output as the sum of value added and intermediates. Then we deflate all nominal values, including gross output, value added, intermediates and wages, to real values. To obtain real values of output, value added and revenue, we deflate the nominal variables using the output deflator supplied by Brandt et al. (2012). Similarly, we deflate the intermediate input with the input deflator supplied by their paper. The construction of real capital stock also follows them. All economic values are therefore in 1998 Chinese *yuan*.

Firm's Age. We use firm's birth year to construct the age of a firm in ASIE. We first clean the reported birth year *bdat*, by adding 1900 to reports which are between 50 and 100, then we replace those who are smaller than 1900. We then compare, by firm, the smallest value in the reported birth year and the earliest year of wave that the firm appears in the sample. If the former is greater than the latter, then we replace the birth year by missing. Instances like this occur when firms go through a slight change in name and the original firm ID, which are nevertheless matched through the panel construction algorithm for sharing the same address, telephone number and legal person for example. The birth year reported is then for the restructured firm, while we have no information when the original firm is established. The age of a firm is then the survey year minus the birth year plus one, so the smallest age is 1. Similarly we construct firm's age using birth year in Industrial Census.

A.2 Estimating Productivity in ASIE

Output elasticities. We use the control function approach of Levinsohn and Petrin (2003) and adapt it to reflect the endogenous growth nature of our theoretical framework. In the benchmark, we estimate value-added Cobb-Douglas production functions at the 2-digit industry level in our sample, using material demand as the proxy for the unobserved persistent productivity shock. To align the estimation closer to the model, we further assume that the persistent productivity shocks follow a random walk with a drift, where the drift is a function of the current period's capital. We allow the state ownership status of the firm to impact both the demand for materials and the random walk process. The construction of the markups relies on the first order condition for labor in the cost minimization problem in the spirit of Hall et al. (1986).[25] We estimate the production functions with GMM in a two-stage estimation procedure.

The value added production function in logs is

$$y_{it} = \beta_0 + \beta_l l_{it} + \beta_k k_{it} + \omega_{it} + \epsilon_{it}, \tag{1}$$

where y, l and k are log value added, log employment and log capital stock respectively and ω is the persistent productivity shock and ϵ is the transitory productivity shock. The timing assumption of the production process is as follows. At the beginning of a period, a firm with capital k_{it} and last period's productivity ω_{it-1} observes the realization of its current period productivity ω_{it}, which depends on k_{it}, ω_{it-1}, and its SOE status. Observing ω_{it}, the firm purchases material, hires labor and produces. Then it makes investment to achieve k_{it+1}, which matters for both improving the productivity next period, generating endogenous growth, and expanding the capital input in production next period.

Maintaining the monotonicity assumption of material demand in productivity as LP, we invert the material demand function to recover the persistent productivity:

$$m_{it} = f(\omega_{it}, k_{it}, z_{it}) \Rightarrow \omega_{it} = h(m_{it}, k_{it}, z_{it}),$$

where z is an indicator for state ownership. Here the inclusion of z allows for the possibility that state-owned firms may face different prices or procure from designated suppliers.

The persistent component of the productivity shock follows a random walk with an en-

[25]For robustness checks, we estimate three variations of the benchmark: (1) from the benchmark, we further include last period's investment or material in the determinant of the drift in the random walk of the productivity; (2) we estimate the benchmark on a "balanced" sample of firms which appear both before 2000 and after 2005; (3) we implement a standard LP but including a time trend in the first stage and modeling the productivity shocks as an AR(1) process, which we view as an "exogenous" growth approach. In (1), we found for a large majority of industries, the impact of last period's material or investment on the drift, after controlling for capital, is close to zero and insignificant. The productivity measure constructed by the estimates from (2) is highly correlated with our benchmark productivity measure at 0.5 and the time trends of the average productivity are almost identical, with the average productivity in 2007, the last year of our sample, being slightly lower at 2.68 under the production parameters estimated off the "balanced panel" than the 2.77 under the benchmark. In (3), average productivity grows slightly less over time from 1 to 2.43 in 2007 than under the benchmark, but the correlation of productivity with the benchmark measure is still high at 0.62.

dogenous drift:

$$\omega_{it} = \omega_{it-1} + \delta_0 + \delta_k k_{it} + \delta_z z_{it} + \delta_h h_{it-1} + \xi_{it}, \tag{2}$$

where the drift term, $\delta_0 + \delta_k k_{it} + \delta_z z_{it} + \delta_h h_{it-1}$, is a deterministic function of current capital, SOE status, and last period's HHI. Here we again allow for the possibility that state ownership impacts the evolution of productivity directly.

In the first stage, we regress y_{it} on a constant, k_{it}, l_{it}, m_{it} and z_{it} assuming a linear functional form for $h(\cdot,\cdot,\cdot)$, and denote the coefficients by γ_0, γ_k, γ_l, γ_m and γ_z respectively. We recover from this stage the true output elasticity of labor, $\gamma_l = \beta_l$. Then we construct the ϕ_{it} from the first stage estimates of coefficients and data,

$$\begin{aligned}\phi_{it} &= \gamma_0 + \gamma_k k_{it} + \gamma_m m_{it} + \gamma_z z_{it} \\ &= \beta_0 + \beta_k k_{it} + \omega_{it}.\end{aligned} \tag{3}$$

In the second stage, for a guess of β_k and δ_k, combining (2) and (3), we obtain

$$\omega_{it} - \omega_{it-1} - \delta_k k_{it} = \phi_{it} - \beta_k k_{it} - (\phi_{it-1} - \beta_k k_{it-1}) - \delta_k k_{it}, \tag{4}$$

and regress onto the SOE status z_{it} to obtain the residuals $\xi(\beta_k, \delta_k)$. Now we can search over all positive β_k and δ_k to evaluate the the moment conditions:

$$\xi_{it}(\beta_k, \delta_k) \begin{bmatrix} l_{it-1} \\ k_{it} \\ m_{it-1} \end{bmatrix} = 0.$$

We estimate the model with GMM and bootstrap the standard errors with 50 replications.

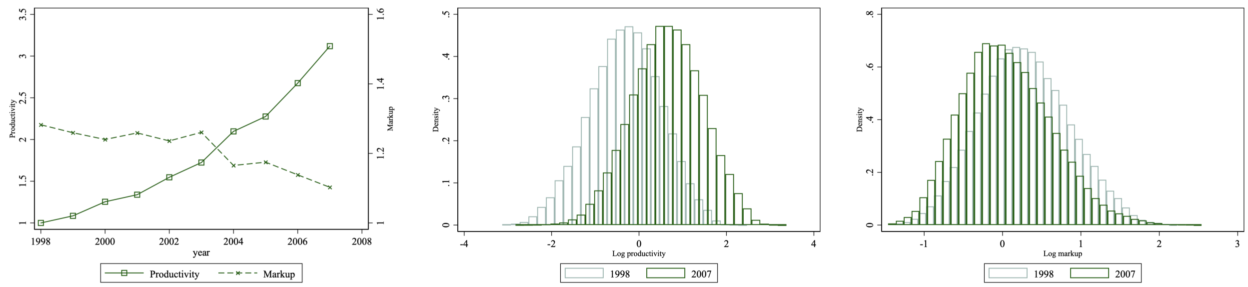
Productivity. To isolate productivity from the markup, we follow the literature and use the cost minimization of the variable labor input to construct the markup $\mu_{it} = \frac{\beta_l}{\alpha_{it}^l}$, where α_{it}^l is the wage bill share of value added: $\alpha_{it}^l = \frac{w_{it} L_{it}}{Y_{it}/exp(\epsilon_{it})}$. Then the firm-level productivity is constructed as

$$exp\left(y_{it} - \widehat{\beta}_l l_{it} - \widehat{\beta}_k k_{it}\right) / \mu_{it}.$$

Because the productivity does not have a natural unit, we normalize the level of the productivity such that the average productivity in any 2-digit industry in 1998, the first year in the sample, is one.

Figure A.1 show the time trend of the estimated productivity and markup in the analysis sample, together with their distributions in 1998 and 2007.

Figure A.1: Estimated Productivity and Markup, ASIE



(a) Productivity (b) Productivity (c) Markup

Note: This figure shows the evolution of average productivity and markup (Panel (a)), the distribution of productivity in 1998 and 2007 (Panel (b)) and the distribution of markup in 1998 and 2007 (Panel (c)).

A.3 Sample Selection

A.3.1 Sample Selection in Industrial Census

We restrict the Census sample to the manufacturing sector and drop observations with missing values of age, employment, ownership type, revenue (called business income in the sample), real revenue, and the industry code. For industry-level analysis, we further drop the industries with less than 100 observations each year. Table A.1 summarizes the sample selection process.

A.3.2 Sample Selection in ASIE

We keep the manufacturing industries, that all 4-digit CIC codes between 1300 and 4400. To prepare the data for the estimation of the production function, we first implement an internal data consistency check in the raw data: 1. We compute an implied wage rate for each firm by dividing its real wage bill by employment and replace the wage bill and employment by missing for firms in the top and bottom 5% of the implied wage rate distribution, by two-digit industry and year. 2. We compute an implied capital rate of return by dividing the capital income, which is the difference between real value added and real wage bill, by the stock of real capital. We then replace the value added, wage bill and

Table A.1: Census Sample Selection

	1995	2004	2008
Raw Data	687,478	1,328,891	1,823,848
Drop missing age	451,797	1,280,683	1,810,000
Drop missing emp	450,532	1,280,683	1,810,000
Drop missing soe	450,532	1,280,683	1,810,000
Drop missing industry code	450,532	1,280,683	1,810,000
Drop missing revenue	424,479	1,280,682	1,810,000
Drop missing real revenue	417,702	1,264,431	1,778,261
Drop if harmonized industry code	417,702	1,264,431	1,778,261
Drop industries with less than 100 obs	414,989	1,263,108	1,777,504

Note: This table reports the change of the number of observations at each step of the sample selection process in the construction of the census sample.

capital stock by missing for firms in the top and bottom 5% of the implied capital return distribution, by two-digit industry and year. 3. We compute an implied labor share share by dividing real wage bill by real value added and replace wage bill and value added by missing for firms who are in the top and bottom 5% of the labor share distribution, by two-digit industry and year. At the end of this check, we drop all firms with missing or negative value added, output, employment, capital stock, material and wage bill. Then we trim the top 1% of these six variables by two-digit industry and year. We further drop firms with missing state ownership indicator or location code. This leads to a panel of firms on which we estimate productivity and markup.

After estimating firms' output elasticities, we drop firms whose labor share (without transitory shock) is above 1, which would lead to unreasonably low markup. We then drop the 2-digit industries whose output elasticities are insignificant at 5% level.[26] We further trim the top and bottom 1% of the level and growth rate of productivity and markup, by two-digit industry and year. Finally we keep all firms with non-missing productivity, markup and age. Table A.2 provides a summary of the sample selection process.

A.4 Supplementary Information of the Figures and Tables

Figure 1.1 In Panel (a), we report the real revenue per worker for the industrial sector as well as the employment-weighted productivity estimated from our ASIE sample. We

[26]More specifically, three industries are dropped: tobacco processing industry (CIC code 16), clothing, shoes and hat manufacturing (CIC code 18), wood processing and bamboo, rattan, palm and grass products industry (CIC code 20), which as a total account for 7.2% of the sample. See Table A.2.

Table A.2: ASIE Sample Selection

	Observations	Deleted obs
Raw Data (manufacturing only)	2,051,162	
Internal data consistency check	1,318,243	732,919
Trim output, inputs and wage bill	1,291,729	26,514
Drop missing ownership type	1,291,722	7
Drop missing location	1,291,678	44
Drop if labor share bigger than 1	1,283,619	8,059
Drop insignificant output elasticities	1,196,863	86,756
Trim level & growth of productivity and markup	1,134,712	62,151
Drop if missing age	1,134,614	98
Final analysis sample	1,134,614	

Note: This table reports the change of the number of observations at each step of the sample selection process in the construction of the ASIE analysis sample.

construct the revenue per worker by dividing the main business income of industrial enterprises by the number of workers in the industrial sector, both obtained from the NSB. The business income dates back to 1979, where our series begin. We deflate the business income by the GDP deflator for China downloaded from the World Bank. To facilitate comparison with the productivity estimates from the firm-level panel, we normalize the level of real revenue per worker to be one in 1998. The share of SOE is calculated from the number of firms by ownership types from the NBS. We classify solely state-owned enterprises, collectively-owned enterprises, joint-stock companies and cooperatives and state-owned limited liability companies as SOEs. The share of SOE is the number of SOEs divided by the total number of industrial firms. This classification of ownership is only adopted by NBS in 1998, therefore we don't have direct evidence of share of SOE before 1998. However it is reasonable to assume that the state ownership is extremely high before 1998, as the privatization of the state sector started only in late 1990s.

In Panel (b), two measures of aggregate entry rate are reported. The entry rates constructed from the 1995 (2004) Industrial Census are calculated as the number of new firms in 1991 (2001) to 1995 (2004) as a fraction of the number of total firms in 1995 (2004). The entry rates from the Business Registry and NBS data are calculated as the number of newly registered firms in the Business Registry Record divided by the total number of active firms reported by the NBS for the industrial sector. The adjustments are made to the raw entry rates constructed from Business Registry Records and NBS, because the NBS made two structural changes in its reporting of the total number of active firms in 1998 and in 2004. To the extent that the Industrial Census is the most accurate data source for the purpose of computing entry rates, we adjust the level of the raw entry rates computed from the Business Registry and NBS over the period 1960-1998 up so that the average entry rate in the 1991-5 period matches that from the 1995 Industrial Census. We apply the same adjustment to 1999-2004 and 2005-2008, benchmarking the levels of those from the Business Registry and NBS data to those in the 2004 and 2008 Industrial Census.

Figure 1.2 In Panel (a), entry rate is defined as the ratio of the average number of firms established in 2003 and 2004 to the total number of incumbent firms in 2004, weighted by employment, in the 2004 Census data. In Panel (b), the revenue growth rate is the annualized growth rate of industry total revenue from 2004 to 2008, using the 2004 and 2008 Census data. In Panel (c), HHI is calculated by industry from the 2008 Census. Industry is defined at the 4-digit CIC level.

Appendix B Additional Empirical Results

B.1 HHI and Revenue Growth by Industry

In Section 2, we have shown that higher entry rates are associated with lower industry-level HHI and higher revenue growth in subsequent years. Table B.1 further shows that empirically there is also a significant and negative correlation between the HHI in 2004 and real revenue growth from 2004 to 2008 across industries, using the 2004 and 2008 Industrial Census data.

Table B.1: Industry-Level HHI and Real Revenue Growth

	(1)	(2)	(3)
log HHI	-0.00906**	-0.0137**	-0.0181***
	(-2.12)	(-2.47)	(-3.37)
number of firms (million)		-1.687	-2.202***
		(-1.32)	(-2.95)
R^2	0.011	0.015	0.139
2-digit industry F.E.	No	No	Yes
2-digit industry clustered S.E.	No	No	Yes
Observations	400	400	400

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: This table shows the results of regressing industry-level annual real revenue growth from 2004 to 2008 on industry-level log HHI in 2004, controlling for number of firms in 2004.

B.2 SOE and Revenue Growth by Industry

In this section, we examine if the broad pattern of declining SOE share and rising revenue growth we document for the whole manufacturing sector also holds at the industry-level. Specifically, we show, at the two-digit industry level, the evolution of share of SOE and revenue per worker from the ASIE panel. Similar to the aggregate time series in Figure 1.1, in almost all industries in the manufacturing sector, SOE shares decline as revenue per worker (in 10,000 *yuan*) grows over the period of 1998-2007 (Figures B.1-B.3). Across industries, some industries have a higher SOE share initially in 1998 and decline at a faster speed than others, but the broad pattern is there in virtually all industries.

More formally, we establish industry-level relation between change in SOE share and change in revenue growth in the following way. Using the three waves of Industrial Census, 1995, 2004 and 2008, we compute at the four-digit industry level, the average annual growth rate of SOE share and revenue from 1995 to 2004 and 1995 to 2008. Then we regress the revenue growth on the SOE share growth over the corresponding period, controlling for 2-digit industry fixed effect and the 4-digit industry's initial conditions such as the number of firms and average employment in 1995. The results are found in Table B.2. The negative correlation between SOE share growth and revenue growth holds for the shorter 1995-2004 period as well as the longer 1995-2008 period.

It is clear from the industry-level evidence that the empirical regularity that we focus on happens across all industries in the manufacturing sector, albeit with different speed. This means the aggregate phenomenon is not driven by a few industries and motivates our modeling assumption of having two sectors with high and low entry cost and otherwise symmetric industries and our interpretation that one sector has achieved the reduction of entry barrier through the privatization earlier (by 2004) than the other sector.

Table B.2: Industry Real Revenue Growth and SOE Share Growth

	1995-2004		1995-2008	
	(1)	(2)	(3)	(4)
SOE growth rate (95-04)	-0.709***	-0.628***		
	(-4.40)	(-3.55)		
SOE growth rate (95-08)			-0.545***	-0.408***
			(-4.40)	(-3.06)
number of firms (million)	-4.189**	-5.762**	-3.062**	-5.038***
	(-2.09)	(-2.47)	(-2.64)	(-3.42)
average employment		-0.0236***		-0.0297***
		(-4.76)		(-7.03)
Constant	0.0261	0.164***	0.111***	0.284***
	(0.78)	(3.86)	(6.11)	(9.94)
R^2	0.205	0.237	0.201	0.280
Observations	358	358	358	358

t statistics in parentheses

2-digit industry F.E. controlled; standard errors clustered at 2-digit industry level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: This table reports the results of regressing industry real revenue growth on industry SOE share growth (column [1]-[2] for periods 1995-2004; column [3]-[4] for periods 1995-2008), controlling for two-digit CIC industry fixed effect and four-digit industry's total number of firms and average employment in 1995. Real revenue and SOE share growth is calculated as annualized growth rates from 1995 to 2004 or from 1995 to 2008 for each four-digit CIC industry.

B.3 The Role of Foreign Firms and Export

Our sample period also covers China's entry into WTO in 2001, which in principle facilitated both foreign firms entering the Chinese market and Chinese firms exporting to other countries. In this section, we examine the role of foreign firms as well as the role of domestic exporting firms.

Foreign owned firms From our Census sample, entrants in 1995 are mostly state and foreign owned, while by 2004 entrants are predominantly privately owned (Table B.3).

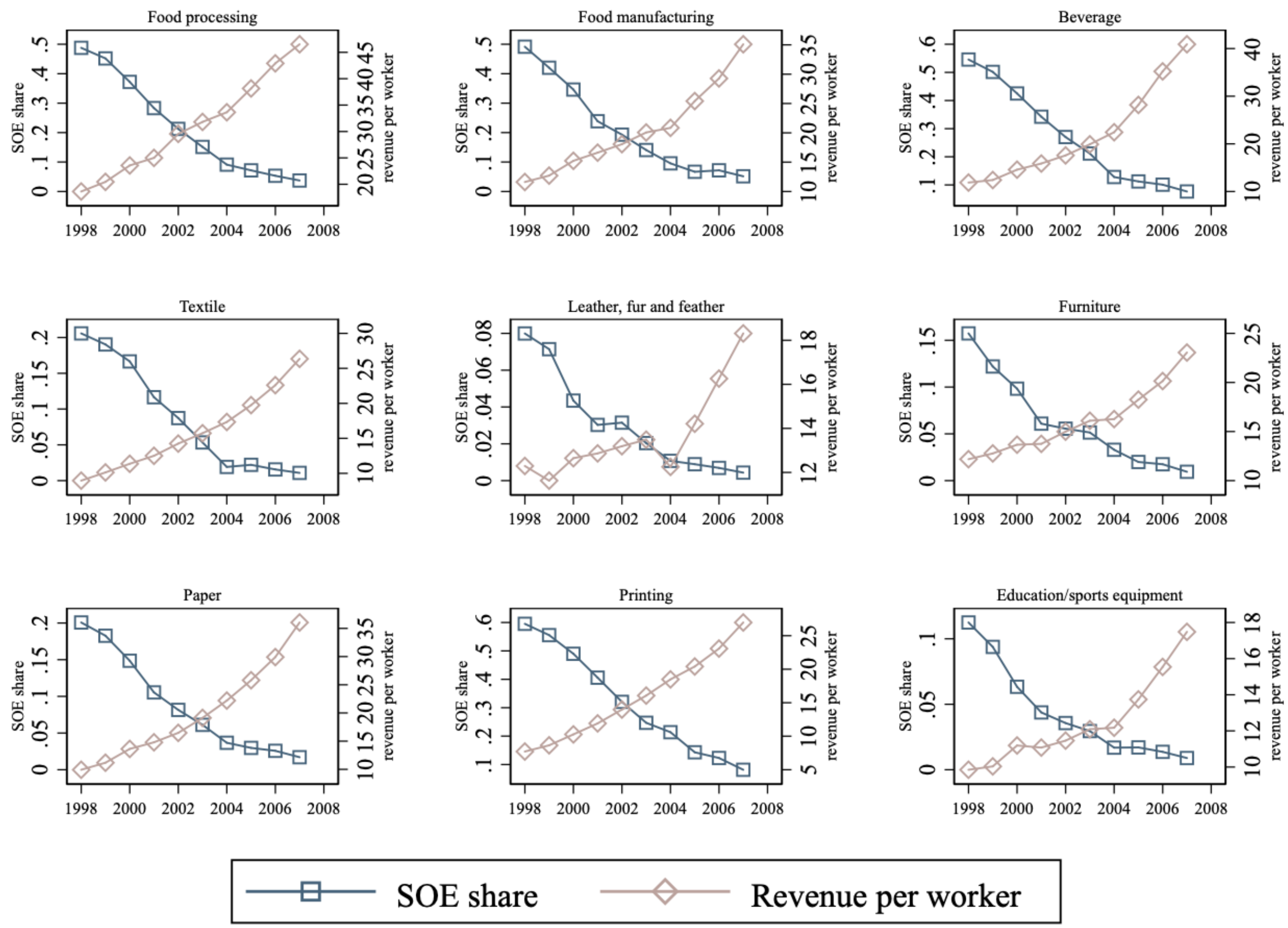


Figure B.1: Evolution of SOE Share and Revenue by Industry, 1998-2007 (I)
Note: This figure plots the evolution of SOE share and revenue per worker by industry.

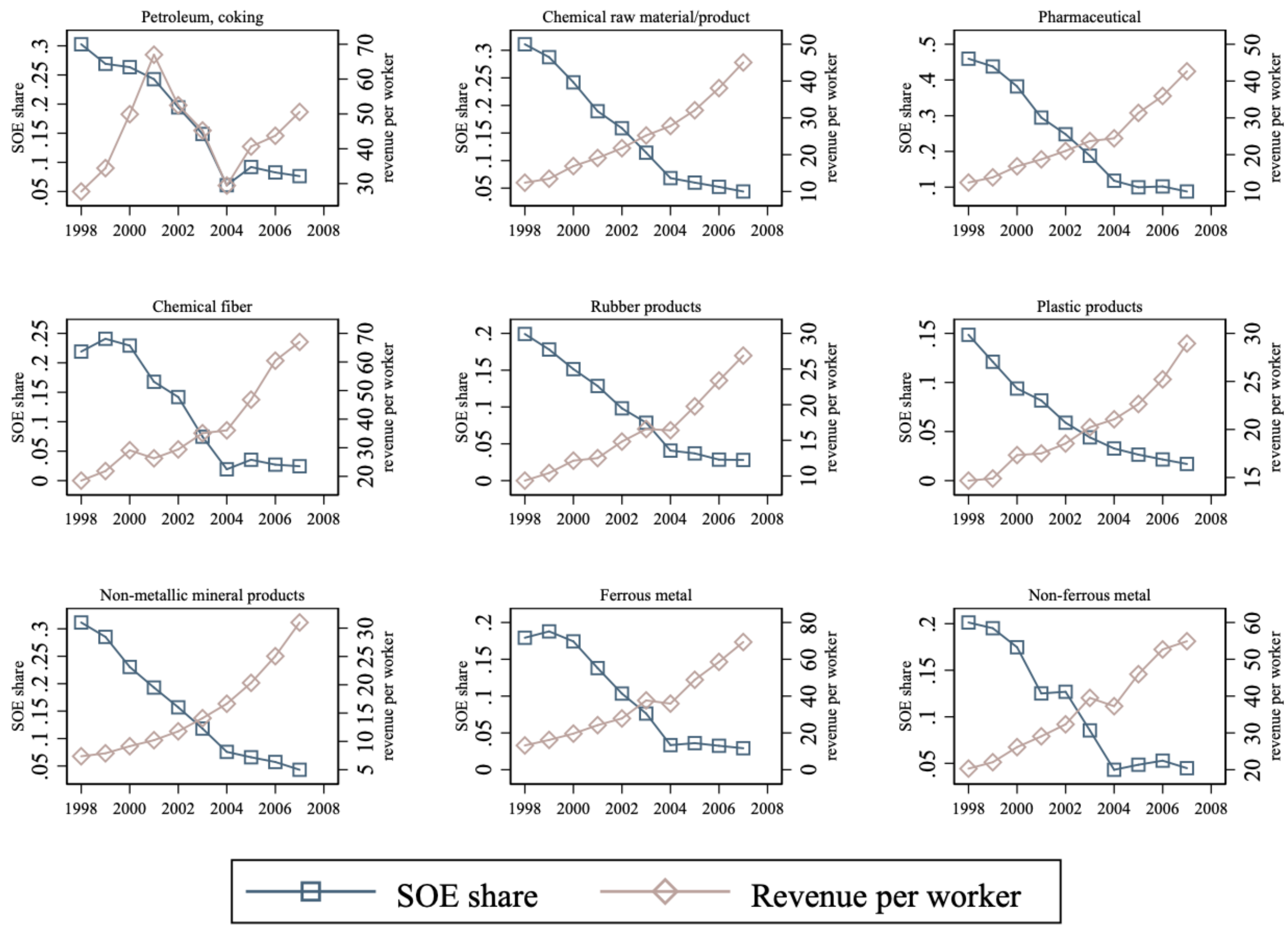


Figure B.2: Evolution of SOE Share and Revenue by Industry, 1998-2007 (II)
Note: This figure plots the evolution of SOE share and revenue per worker by industry.

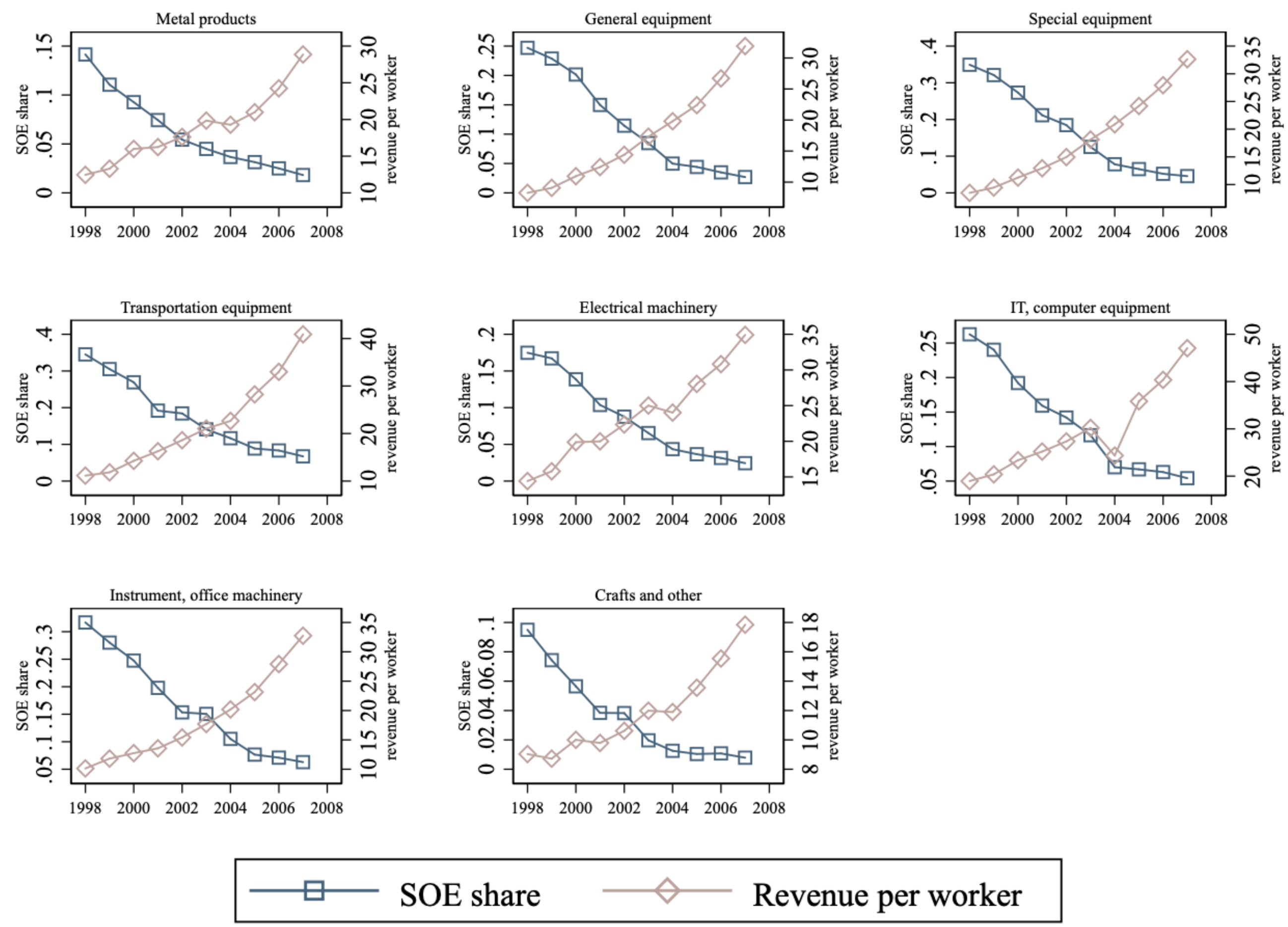


Figure B.3: Evolution of SOE Share and Revenue by Industry, 1998-2007 (III)

Note: This figure plots the evolution of SOE share and revenue per worker by industry.

An interesting question is whether FOEs behaved differently than POEs or whether pro-competitive and growth-enhancing effects could be driven by the entry of FOEs instead of domestic POEs. If so, the framework to examine the pro-competitive and pro-growth effect would necessarily entail an international trade element. We provide evidence against this hypothesis.

To do that, we rerun the empirical specifications similar to that in Table 2.2 but distinguishing the entry rates of SOEs, POEs and FOEs separately. Specifically, we regress the four-digit industry level annualized revenue growth from 2004 to 2008 and log HHI in 2008 on the entry rates of three types of enterprises (SOE, POE and FOE) in 2004, controlling for 2-digit industry fixed effects and 4-digit industry characteristics in 2004. The results are in Table B.4. Clearly, the empirical pattern that we focus on, namely entry being positively correlated with growth and negatively correlated with market power, is driven by domestic POE entry.

Table B.3: Employment Share of New Entrants by Ownership Types (%)

	1995	2004	2008
SOE share	65.12	3.61	3.33
POE share	5.64	77.59	83.06
FOE share	29.25	18.80	13.60

Note: This table shows the employment shares of entrants by ownership types in 1995, 2005 and 2008 Census.

Table B.4: Industry-Level Entry Rate, Real Revenue Growth and HHI by Ownership Types

	Real Revenue Growth			log HHI		
	(1)	(2)	(3)	(4)	(5)	(6)
SOE entry	2.796	2.976	2.844	15.45	14.53	17.54
	(0.94)	(1.01)	(1.01)	(1.13)	(1.02)	(1.18)
POE entry	0.862***	0.466**	0.417**	-7.091***	-5.070**	-3.947*
	(4.86)	(2.65)	(2.59)	(-3.73)	(-2.64)	(-1.97)
FOE entry	0.123	0.341	0.377	-3.957	-5.071	-5.883*
	(0.45)	(1.40)	(1.63)	(-1.26)	(-1.62)	(-1.87)
number of firms (million)	-0.0466	-0.875	-1.023	-123.6***	-119.3***	-116.0***
	(-0.07)	(-1.27)	(-1.43)	(-6.28)	(-6.41)	(-6.15)
log employment		-0.0469***	-0.0246		0.239**	-0.267*
		(-5.82)	(-1.38)		(2.74)	(-1.90)
log revenue			-0.0194			0.441***
			(-1.51)			(4.09)
R^2	0.196	0.245	0.251	0.521	0.531	0.552
Observations	400	400	400	400	400	400

t statistics in parentheses

2-digit CIC industry fixed effects controlled; standard errors clustered at 2-digit industry level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: This table reports the correlation between industry-level real revenue growth (column [1]-[3]) and concentration measured by HHI (column [4]-[6]) and entries with various ownership types. Entry rate is defined as the ratio of the average number of firms established in 2003 and 2004 to the total number of incumbent firms in 2004, weighted by employment for each ownership type. Total real revenue is calculated by industry in 2004 and 2008 and growth rate is calculated as the average annual growth rate within the time period. HHI index is calculated by industry in 2008. Industry is defined at 4-digit CIC level.

Exporting firms On the other hand, the entry to WTO allows more domestic firms to export and we ask if entry could be driven by the reduction in trade barrier to export. To see that, we examine if in industries with more entry, we see more export. The evidence does not seem to support this view. Figure B.4 shows two scatter plots using the ASIE panel.[27] In this figure, we plot both the four-digit industry-level export to revenue ratio averaged from 2005-7 and the annualized growth rate of the export to revenue ratio from 2005 to 2007 against the industry's entry rate in 2004. Neither the level nor the growth of export to revenue ratios correlate with entry in 2004 across industries.

Even though entry does not systematically correlate with export, one might still ask if the productivity growth mostly accrues to firms who export. We then use firm-level productivity growth constructed from ASIE and regress it on a dummy of firms who export (column [1]), on a dummy of firms who export more than 50% of output conditioning on exporting at all (column [2]), and on the share of output that is exported and the share squared (column [3]), controlling for industry-year-province fixed effects. We find an inverted-U relationship between firm's exporting behavior and its productivity growth. That is, exporting firms on average do achieve higher growth than non-exporting firms, but conditioning on exporting, export-oriented firms whose main market is foreign achieve a lower growth than exporting firms whose main market is domestic. We interpret these results as evidence that even though export may contribute to productivity growth, productivity growth is not driven by exports and firms who tend to achieve higher productivity growth are still those facing mainly a domestic market, whose market structure has changed over time.

[27] Census data do not contain export values. ASIE has export values in all years except 2004.

Table B.5: Productivity Growth of Exporting Firms, ASIE

	Productivity growth		
	(1)	(2)	(3)
Positive share of export	0.0272***		
	(11.37)		
Export more than 50% of output		-0.0180***	
		(-5.43)	
Share of export			0.149***
			(11.27)
Share of export squared			-0.152***
			(-10.82)
R^2	0.117	0.146	0.117
Observations	566,559	140,439	566,559

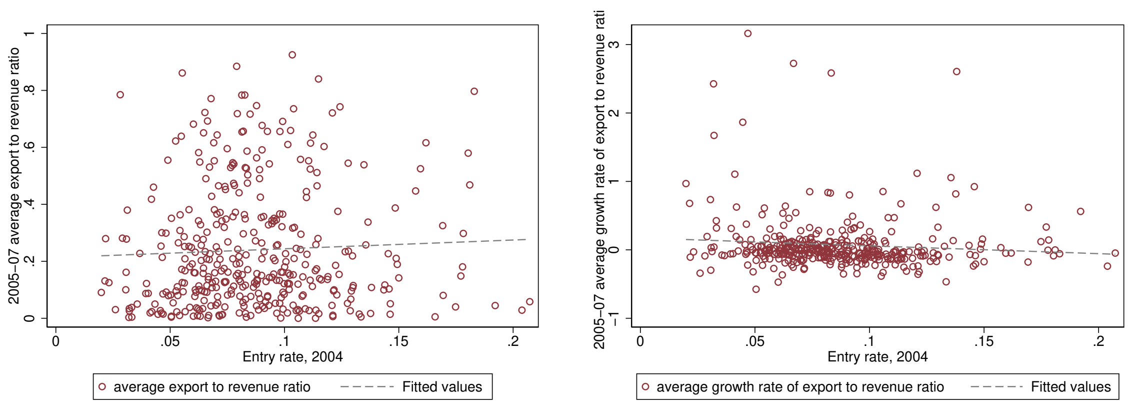
t statistics in parentheses

4-digit-industry-year-province fixed effects controlled; standard errors clustered at 4-digit industry level

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: The table reports the results from regressing productivity growth on a dummy of exporter (column [1]), on a dummy of exporter which exports more than 50% of output (column [2]) and on share of export and its square (column [3]), controlling for industry-year-province fixed effects.

Figure B.4: Entry Rate versus Level and Growth of Export Across Industries



(a) average export to revenue ratio, 2005-7 (b) growth of export to revenue ratio, 2005-7

Note: This figure shows the scatter plots between 2004 industry-level entry rates and 2005-2007 average export to revenue ratio or 2005-2007 average growth rates of export to revenue ratio. Entry rate is defined as the ratio of the average number of firms established in 2003 and 2004 and the total number of incumbent firms in 2004, weighted by employment from 2004 Census. Export to revenue ratio is defined as the ratio of average industry export and average industry real revenue in ASIE.

Appendix C Model and Calibration

Inflow and Outflow Tables In this section, we present the inflow to and outflow from an industry with a given n for the other three type categories of the leader-follower pair: (h,l), (l,h) and (l,l), in Tables C.1-C.3. The table for the (h,h) configuration is Table 3.1 in the paper.

Table C.1: Inflow and Outflow in Industry $(i = h, j = l)$

State	*Inflow*		*Outflow*
n=0:	$\sum_{n\geq 2}\mu_{hl}(n)\overline{x}_{hl}(n)\phi + \sum_{n\geq 2}\mu_{lh}(n)[\overline{x}_{lh}(n)+x^e_{lh}(n)]\phi +$ $\sum_{n\geq 2}\mu_{ll}(n)x^e_{ll}(n)\phi + 2\mu_{hh}(0)\sigma +$ $\mu_{hl}(1)\overline{x}_{hl}(1)+\mu_{lh}(1)[\overline{x}_{lh}(1)+x^e_{lh}(1)]+\mu_{ll}(1)x^e_{ll}(1)$	=	$\mu_{hl}(0) * [x^h_{hl}(0)+x^l_{hl}(0)+x^e_{hl}(0)+\sigma]$
n=1:	$\mu_{hl}(0)[x^h_{hl}(0)+x^e_{hl}(0)/2]+\mu_{ll}(0)x^e_{ll}(0)+\mu_{hh}(1)\sigma +$ $\mu_{hl}(2)\overline{x}_{hl}(2)(1-\phi)$	=	$\mu_{hl}(1)[x_{hl}(1)+\overline{x}_{hl}(1)+x^e_{hl}(1)+\sigma]$
n≥ 2 :	$\mu_{hl}(n-1)x_{hl}(n-1)+\mu_{hh}(n)\sigma +$ $\mu_{hl}(n+1)\overline{x}_{hl}(n+1)(1-\phi)$	=	$\mu_{hl}(n)[x_{hl}(n)+\overline{x}_{hl}(n)+x^e_{hl}(n)+\sigma]$

Note: This table lists the inflow to and outflow from all possible states (i.e. gap sizes) given a (h,l) leader-follower configuration.

Table C.2: Inflow and Outflow in Industry $(i = l, j = h)$

State	*Inflow*		*Outflow*
n=0:	$\sum_{n\geq 2}\mu_{hl}(n)\overline{x}_{hl}(n)\phi + \sum_{n\geq 2}\mu_{lh}(n)[\overline{x}_{lh}(n)+x^e_{lh}(n)]\phi +$ $\sum_{n\geq 2}\mu_{ll}(n)x^e_{ll}(n)\phi + 2\mu_{hh}(0)\sigma +$ $\mu_{hl}(1)\overline{x}_{hl}(1)+\mu_{lh}(1)[\overline{x}_{lh}(1)+x^e_{lh}(1)]+\mu_{ll}(1)x^e_{ll}(1)$	=	$\mu_{hl}(0)[x^h_{hl}(0)+x^l_{lh}(0)+x^e_{hl}(0)+\sigma]$
n= 1:	$\mu_{hl}(0)x^l_{hl}(0)+\mu_{hh}(1)\sigma+\mu_{lh}(2)[\overline{x}_{lh}(2)+x^e_{lh}(2)](1-\phi) +$ $\mu_{ll}(2)x^e_{ll}(2)(1-\phi)$	=	$\mu_{lh}(1)[x_{lh}(1)+\overline{x}_{lh}(1)+x^e_{lh}(1)+\sigma]$
n≥ 2 :	$\mu_{lh}(n-1)x_{lh}(n-1)+\mu_{hh}(n)\sigma+\mu_{ll}(n+1)x^e_{ll}(n+1)(1-\phi) +$ $\mu_{lh}(n+1)[\overline{x}_{lh}(n+1)+x^e_{lh}(n+1)](1-\phi)$	=	$\mu_{lh}(n)[x_{lh}(n)+\overline{x}_{lh}(n)+x^e_{lh}(n)+\sigma]$

Note: This table lists the inflow to and outflow from all possible states (i.e. gap sizes) given a (l,h) leader-follower configuration.

Table C.3: Inflow and Outflow in Industry $(i = l, j = l)$

State	*Inflow*		*Outflow*
n=0:	$\mu_{ll}(1)\overline{x}_{ll}(1) + \sum_{n\geq 2} \mu_{ll}(n)\overline{x}_{ll}(n)\phi + \mu_{hl}(0)\sigma$	=	$\mu_{ll}(0)[2x^l_{ll}(0) + x^e_{ll}(0)]$
n=1:	$2\mu_{ll}(0)x^l_{ll}(0) + [\mu_{hl}(1) + \mu_{lh}(1)]\sigma + \mu_{ll}(2)\overline{x}_{ll}(2)(1-\phi)$	=	$\mu_{ll}(1)[x_{ll}(1) + \overline{x}_{ll}(1) + x^e_{ll}(1)]$
n$\geq$ 2 :	$\mu_{ll}(n-1)x_{ll}(n-1) + [\mu_{hl}(n) + \mu_{lh}(n)]\sigma +$ $\mu_{ll}(n+1)\overline{x}_{ll}(n+1)(1-\phi)$	=	$\mu_{ll}(n)[x_{ll}(n) + \overline{x}_{ll}(n) + x^e_{ll}(n)]$

Note: This table lists the inflow to and outflow from all possible states (i.e. gap sizes) given a (l, l) leader-follower configuration.

Aggregate output growth Define $\ln Y_1(t) \equiv \int_0^\zeta \ln y_\nu(t)d\nu$ and $\ln Y_2(t) \equiv \int_\zeta^1 \ln y_\nu(t)d\nu$. Further define

$$g \equiv \frac{d\ln Y(t)}{dt} = \int_0^1 \frac{d\ln y_\nu(t)}{dt}d\nu,$$

$$g_1 \equiv \frac{d\ln Y_1(t)}{dt} = \int_0^\zeta \frac{d\ln y_\nu(t)}{dt}d\nu,$$

and

$$g_2 \equiv \frac{d\ln Y_2(t)}{dt} = \int_\zeta^1 \frac{d\ln y_\nu(t)}{dt}d\nu.$$

It is straightforward to see that

$$g = g_1 + g_2.$$

We focus here on sector 1, as the analysis for sector 2 is analogous. For any industry, from the industrial production function we have

$$\begin{aligned} y = [y_1^\delta + y_2^\delta]^{\frac{1}{\delta}} &= [p_1^{\delta/(\delta-1)} + p_2^{\delta/(\delta-1)}]^{\frac{1-\delta}{\delta}} \\ &= [(\frac{1-\delta\omega_1}{\delta(1-\omega_1)}c_1)^{\delta/(\delta-1)} + (\frac{1-\delta\omega_2}{\delta(1-\omega_2)}c_2)^{\delta/(\delta-1)}]^{\frac{1-\delta}{\delta}} \\ &= c_2{}^{-1}[(\frac{1-\delta\omega_1}{\delta(1-\omega_1)}\frac{c_1}{c_2})^{\delta/(\delta-1)} + (\frac{1-\delta\omega_2}{\delta(1-\omega_2)})^{\delta/(\delta-1)}]^{\frac{1-\delta}{\delta}} \\ &= \lambda^{n_2}[f_1(n)^{\delta/(\delta-1)} + f_2(n)^{\delta/(\delta-1)}]^{\frac{1-\delta}{\delta}} \end{aligned}$$

where the last equality holds as both ω_1 and ω_2 are determined by the gap n.

Changes in y from a successful innovation by the leader and follower in an industry with gap n, denoted as a_n^L and a_n^F respectively, are

$$\begin{aligned}a_n^L &\equiv \ln \tilde{y} - \ln y \\ &= [f_1(n+1)^{\delta/(\delta-1)} + f_2(n+1)^{\delta/(\delta-1)}]^{\frac{1-\delta}{\delta}} - [f_1(n)^{\delta/(\delta-1)} + f_2(n)^{\delta/(\delta-1)}]^{\frac{1-\delta}{\delta}},\end{aligned}$$

if a follower improves n steps

$$\begin{aligned}a_n^F(n) &\equiv \ln \tilde{y} - \ln y \\ &= n \ln \lambda + [f_1(0)^{\delta/(\delta-1)} + f_2(0)^{\delta/(\delta-1)}]^{\frac{1-\delta}{\delta}} - [f_1(n)^{\delta/(\delta-1)} + f_2(n)^{\delta/(\delta-1)}]^{\frac{1-\delta}{\delta}}\end{aligned}$$

if a follower improves 1 step

$$\begin{aligned}a_n^F(1) &\equiv \ln \tilde{y} - \ln y \\ &= \ln \lambda + [f_1(n-1)^{\delta/(\delta-1)} + f_2(n-1)^{\delta/(\delta-1)}]^{\frac{1-\delta}{\delta}} - [f_1(n)^{\delta/(\delta-1)} + f_2(n)^{\delta/(\delta-1)}]^{\frac{1-\delta}{\delta}}\end{aligned}$$

It follows that $\sum_{m=1}^{n} a_{m-1}^L + a_n^F(n) = n \ln \lambda$, and $a_n^L + a_{n+1}^F(1) = \ln \lambda$.

The aggregate growth rate of output in sector 1 is,

$$\begin{aligned}g_1 = &\sum_{i=h,l} \sum_{j=h,l} \sum_{n\geq 1} \mu_{ij}(n)\{x_{ij}(n)a_n^L + (\overline{x}_{ij}(n) + x_{ij}^e(n))[\phi a_n^F(n) + (1-\phi)a_n^F(1)]\} \\ &+ \mu_{hh}(0)(2x_{hh}^h(0) + x_{hh}^e(0))a_0^L + \mu_{hl}(0)(x_{hl}^h(0) + x_{hl}^l(0) + x_{hl}^e(0))a_0^L + \mu_{ll}(0)(2x_{ll}^l(0) + x_{ll}^e(0))a_0^L \\ = &\sum_{i=h,l} \sum_{j=h,l} \sum_{n\geq 1} \mu_{ij}(n)\{x_{ij}(n)a_n^L + (\overline{x}_{ij}(n) + x_{ij}^e(n))[\phi(n \ln \lambda - \sum_{m=1}^{n} a_{m-1}^L) + (1-\phi)(\ln \lambda - a_{n-1}^L)]\} \\ &+ \mu_{hh}(0)(2x_{hh}^h(0) + x_{hh}^e(0))a_0^L + \mu_{hl}(0)(x_{hl}^h(0) + x_{hl}^l(0) + x_{hl}^e(0))a_0^L + \mu_{ll}(0)(2x_{ll}^l(0) + x_{ll}^e(0))a_0^L \\ = &\sum_{i=h,l} \sum_{j=h,l} \sum_{n\geq 1} \mu_{ij}(n)(\overline{x}_{ij}(n) + x_{ij}^e(n))[\phi n \ln \lambda + (1-\phi)\ln \lambda] \\ &+ \sum_{i=h,l} \sum_{j=h,l} \sum_{n\geq 1} \mu_{ij}(n)[x_{ij}(n)a_n^L - (\overline{x}_{ij}(n) + x_{ij}^e(n))(\phi \sum_{m=1}^{n} a_{m-1}^L + (1-\phi)a_{n-1}^L)] \\ &+ \mu_{hh}(0)(2x_{hh}^h(0) + x_{hh}^e(0))a_0^L + \mu_{hl}(0)(x_{hl}^h(0) + x_{hl}^l(0) + x_{hl}^e(0))a_0^L + \mu_{ll}(0)(2x_{ll}^l(0) + x_{ll}^e(0))a_0^L \\ = &\sum_{i=h,l} \sum_{j=h,l} \sum_{n\geq 1} \mu_{ij}(n)(\overline{x}_{ij}(n) + x_{ij}^e(n))[\phi n \ln \lambda + (1-\phi)\ln \lambda]\end{aligned}$$

The second equation follows from $\sum_{m=1}^{n} a_{m-1}^L + a_n^F(n) = n \ln \lambda$, and $a_n^L + a_{n+1}^F(1) = \ln \lambda$. The last equality holds as the rest of terms is equal to zero. To see this point, note that the coefficient in front of a_0^L is

$$-\sum_{i=h,l}\sum_{j=h,l}\sum_{n\geq 1}\mu_{ij}(n)(\overline{x}_{ij}(n)+x^e_{ij}(n))\phi-\sum_{i=h,l}\sum_{j=h,l}\mu_{ij}(1)(\overline{x}_{ij}(1)+x^e_{ij}(1))(1-\phi)$$
$$+\mu_{hh}(0)(2x^h_{hh}(0)+x^e_{hh}(0))+\mu_{hl}(0)(x^h_{hl}(0)+x^l_{hl}(0)+x^e_{hl}(0))+\mu_{ll}(0)(2x^l_{ll}(0)+x^e_{ll}(0))$$

which equals zero under the stationary distribution.

The coefficient in front of $a_n^L, n\geq 1$, is

$$\sum_{i=h,l}\sum_{j=h,l}\mu_{ij}(n)x_{ij}(n)-\sum_{i=h,l}\sum_{j=h,l}\sum_{m\geq n+1}\mu_{ij}(m)\phi(\overline{x}_{ij}(m)+x^e_{ij}(m))$$
$$-\sum_{i=h,l}\sum_{j=h,l}\mu_{ij}(n+1)(1-\phi)(\overline{x}_{ij}(n+1)+x^e_{ij}(n+1))$$

which also equals zero under stationary distribution.

Therefore sector growth

$$g_1=\sum_{i=h,l}\sum_{j=h,l}\sum_{n\geq 1}\mu_{ij}(n)(\overline{x}_{ij}(n)+x^e_{ij}(n))[\phi n\ln\lambda+(1-\phi)\ln\lambda].$$

Note that

$$\sum_{i=h,l}\sum_{j=h,l}\sum_{n=1}\mu_{ij}(n)(\overline{x}_{ij}(n)+x^e_{ij}(n))(n\phi+1-\phi)$$
$$=\sum_{i=h,l}\sum_{j=h,l}\sum_{n=1}\mu_{ij}(n)(\overline{x}_{ij}(n)+x^e_{ij}(n))(1-\phi)+\sum_{i=h,l}\sum_{j=h,l}\sum_{n=1}\mu_{ij}(n)(\overline{x}_{ij}(n)+x^e_{ij}(n))\phi$$
$$+\sum_{i=h,l}\sum_{j=h,l}\sum_{n=2}\mu_{ij}(n)(\overline{x}_{ij}(n)+x^e_{ij}(n))\phi+\sum_{i=h,l}\sum_{j=h,l}\sum_{n=3}\mu_{ij}(n)(\overline{x}_{ij}(n)+x^e_{ij}(n))\phi+...$$
$$=\mu_{hh}(0)(2x^h_{hh}(0)+x^e_{hh}(0))+\mu_{hl}(0)(x^h_{hl}(0)+x^l_{hl}(0)+x^e_{hl}(0))+\mu_{ll}(0)(2x^l_{ll}(0)+x^e_{ll}(0))$$
$$+\sum_{i=h,l}\sum_{j=h,l}\sum_{n=1}\mu_{ij}(n)x_{ij}(n)$$

Therefore

$$g_1\equiv\frac{d\ln Y_1}{dt}=[\sum_{i=h,l}\sum_{j=h,l}\sum_{n\geq 1}\mu_{ij}(n)x_{ij}(n)+\mu(0)x(0)]*\ln\lambda$$

with $\mu(0)x(0)\equiv\sum_{i=h,l}\sum_{i=h,l}\mu_{ii}(0)(2x^i_{ii}(0)+x^e_{ii}(0))+\mu_{hl}(0)(x^h_{hl}(0)+x^l_{hl}(0)+x^e_{hl}(0))$.

Symmetrically we can obtain the growth rate in sector 2, g_2. Note that the average growth rate of an industry in sector 1 is g_1/ζ, and that of an industry in sector 2 is $g_2/(1-\zeta)$. The aggregate growth rate is $g = g_1 + g_2$.

Decomposition of changes in the aggregate growth rate Note that, from the growth rate formula, changes in the growth rate can come from both the x terms, which are summarizes as the direct effect and the Schumpeterian effect, and the μ terms, which can be further decomposed into a replacement effect and a pro-competitive effect. For this later part, there are two symmetric methods.

Method A:

$$\begin{aligned}
&\sum_{\psi}\sum_{n=0} x_1(\psi,n)[\mu_2(\psi,n)-\mu_1(\psi,n)] \\
&\approx \sum_{\psi}\sum_{n=0} x_1(\psi,n)[\tilde{\mu}_2(n)f_2(\psi|n)-\tilde{\mu}_1(n)f_1(\tau|n)] \\
&\approx \sum_{\psi}\sum_{n=0} x_1(\psi,n)[\tilde{\mu}_2(n)f_2(\psi|n)-\tilde{\mu}_1(n)f_2(\psi|n)+\tilde{\mu}_1(n)f_2(\psi|n)-\tilde{\mu}_1(n)f_1(\psi|n)] \\
&= \underbrace{\sum_{\psi}\sum_{n=0} x_1(\psi,n)[f_2(\psi|n)-f_1(\psi|n)]\tilde{\mu}_1(n)}_{\textit{replacement effect}} + \underbrace{\sum_{\psi}\sum_{n=0} x_1(\psi,n)[\tilde{\mu}_2(n)-\tilde{\mu}_1(n)]f_2(\psi|n)}_{\textit{pro-competitive effect}}.
\end{aligned}$$

where $f_s(\psi|n), s = 1,2$ denotes the distribution of ψ conditional on a given value of n, and $\tilde{\mu}_s(n) \equiv \sum_{\psi} \mu_s(\psi,n)$ is the marginal distribution of n in sector $s, s = 1,2$.

Method B:

$$\begin{aligned}
&\sum_{\psi}\sum_{n=0} x_1(\psi,n)[\mu_2(\psi,n)-\mu_1(\psi,n)] \\
&\approx \sum_{\psi}\sum_{n=0} x_1(\psi,n)[\tilde{\mu}_2(\psi)h_2(n|\psi)-\tilde{\mu}_1(\psi)h_1(n|\psi)] \\
&\approx \sum_{\psi}\sum_{n=0} x_1(\psi,n)[\tilde{\mu}_2(\psi)h_2(n|\psi)-\tilde{\mu}_2(\psi)h_1(n|\psi)+\tilde{\mu}_2(\psi)h_1(n|\psi)-\tilde{\mu}_1(\psi)h_1(n|\psi)] \\
&= \underbrace{\sum_{\psi}\sum_{n=0} x_1(\psi,n)[\tilde{\mu}_2(\psi)-\tilde{\mu}_1(\psi)]h_1(n|\psi)}_{\textit{replacement effect}} + \underbrace{\sum_{\psi}\sum_{n=0} x_1(\psi,n)[h_2(n|\psi)-h_1(n|\psi)]\tilde{\mu}_2(\psi)}_{\textit{pro-competitive effect}}.
\end{aligned}$$

where $h_s(n|\psi), s = 1,2$ denotes the distribution of n conditional on a given value of ψ, and $\tilde{\mu}_s(\psi) \equiv \sum_n \mu_s(\psi, n)$ is the marginal distribution of ψ in sector $s, s = 1,2$.

Under the calibrated parameters, Table C.4 presents the decomposition results using these two methods

Table C.4: Decomposition with Method A & B

	replacement effect	pro-competitive effect
Method A	0.0023	0.0190
Method B	0.0151	0.0062
Average	0.0087	0.0126
Average (%)	40.85%	59.15%

Note: This table list the relative contribution of the replacement effect and the pro-competitive effect following two approaches of decomposition.

Method A tends to obtain a larger 'pro-competitive effect', while Method B obtains a larger 'replacement effect'. As there is no clear criteria one method is superior than the other, we use the average from those two methods as the benchmark values for these two effects.